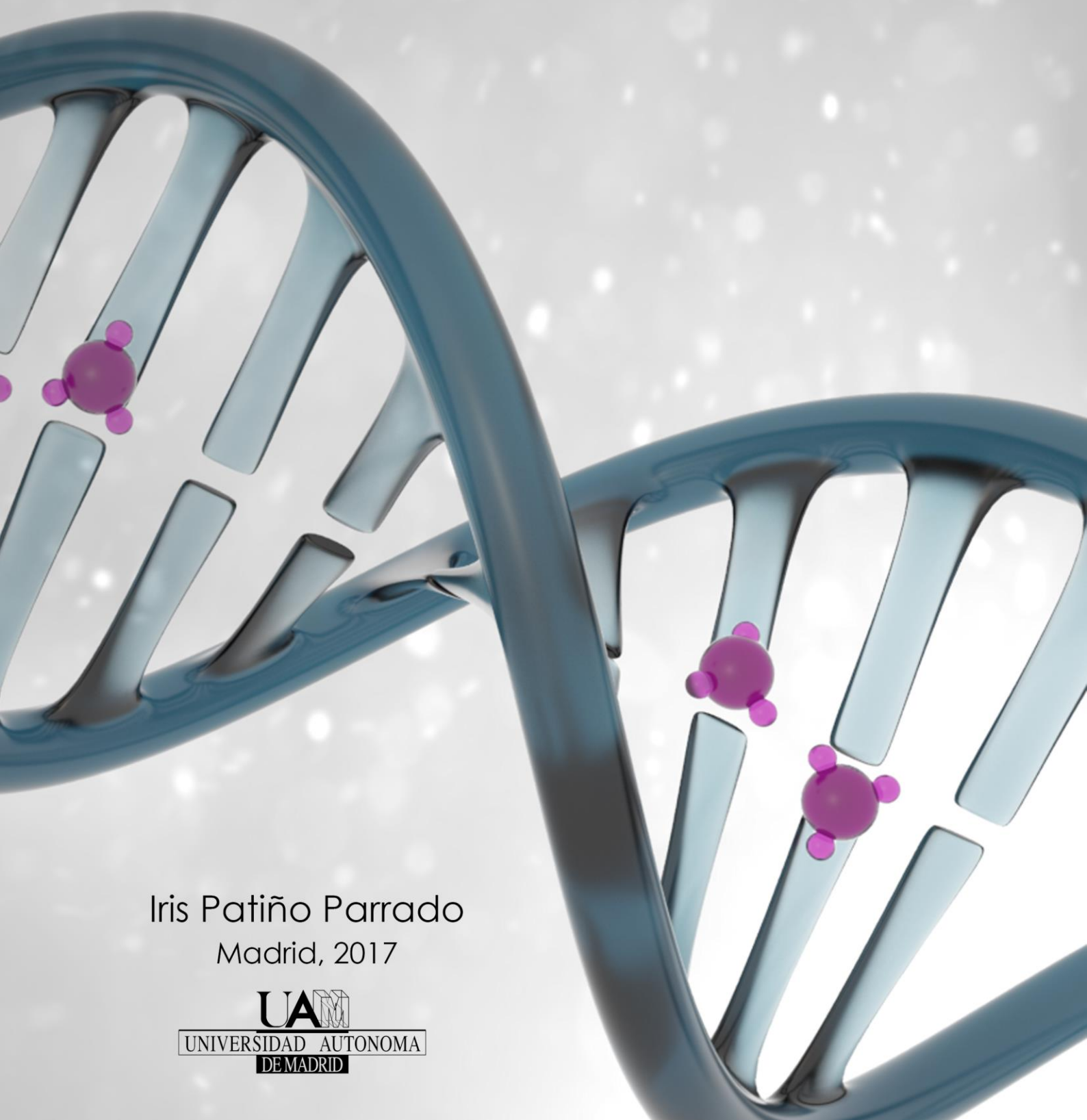


Universidad Autónoma de Madrid
Programa de Doctorado en Biociencias Moleculares

Caracterización del estado de metilación en citosinas en las neuronas tetraploides de la corteza cerebral murina

Iris Patiño Parrado
Madrid, 2017



DEPARTAMENTO DE BIOQUÍMICA
FACULTAD DE MEDICINA
UNIVERSIDAD AUTÓNOMA DE MADRID

**CARACTERIZACIÓN DEL ESTADO DE METILACIÓN EN
CITOSINAS EN LAS NEURONAS TETRAPLOIDES DE LA
CORTEZA CEREBRAL MURINA**

Tesis Doctoral

IRIS PATIÑO PARRADO

Licenciada en Biología

Director

DR. JOSÉ MARÍA FRADE LÓPEZ

Instituto Cajal (CSIC)

Dr. José María Frade López, Investigador Científico del Instituto Cajal (CSIC)

CERTIFICA:

Que la presente tesis doctoral, que lleva por título **Caracterización del estado de metilación en citosinas en las neuronas tetraploides de la corteza cerebral murina**, que presenta Doña Iris Patiño Parrado para optar al grado de Doctor por la Universidad Autónoma de Madrid ha sido realizada bajo mi dirección en el Insituto Cajal (CSIC) y que cumple todos los requisitos para su defensa pública, reuniendo, a mi juicio el suficiente rigo científico para optar al grado de doctor.

Madrid, a 16 de octubre de 2017

Fdo.: Dr. Jose María Frade López

Agradecimientos

En primer lugar, me gustaría dar las gracias al Doctor Jose María Frade López por darme la oportunidad de realizar la tesis en su laboratorio y de formarme como científica.

Gracias a mis compañeros de laboratorio, en especial a Chaska, por darme perspectiva, y a Maria José, por ser como es.

Gracias a Álvaro por dejarme en herencia este tema que finalmente ha dado lugar a una tesis.

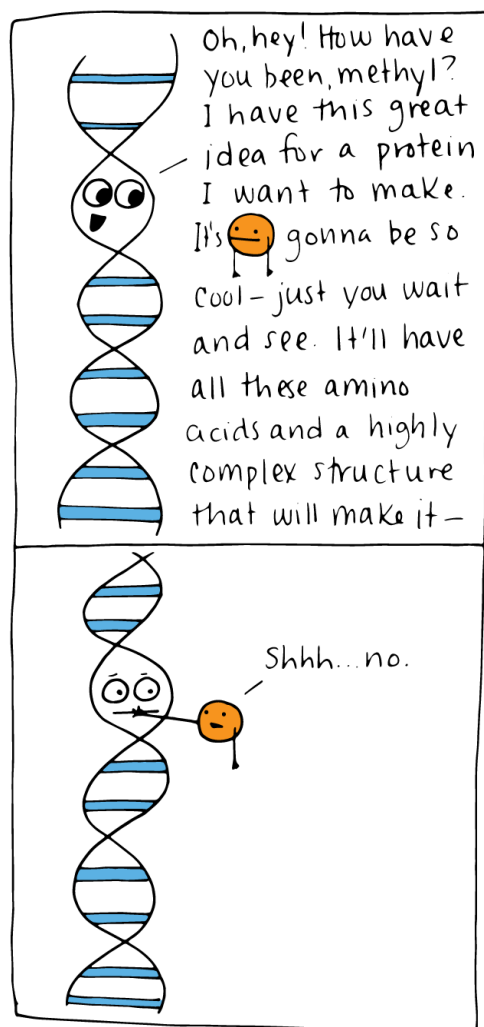
Gracias a Noelia Villarrubia y a Jose Luis por las horas compartidas en el citómetro y a Silvia, por ayudarme siempre. A Yago y María José por hacer mucho más agradable la burocracia.

Gracias al laboratorio de la Doctora Aixa Morales y del Doctor Sergio Casas por abrirme las puertas cuando me perseguían las contaminaciones. Gracias a este último por el apoyo científico y moral.

Infinitas gracias a Mark Strunk por sus consejos sobre pirosecuenciación, por permitirme conocer la técnica de primera mano y por las veces que ha dado prioridad a mis muestras para tener resultados antes.

Thank you Norman and Elissa Roser for making me feel at home when I was in Houston and Kuanwei Shen for helping me at the lab.

Gracias Marcos Fernández Callejo por toda su ayuda con el GemBS. Thank you Angelika and Marjorie for your help with bioinformatic issues.



Another gene silenced.
•Beatrice the Biologist

Agradecimientos

Gracias a mis amigos de toda la vida por seguir a mi lado y compartir esta etapa también. Gracias a mis amigos de la universidad, tesisandos también, por compartir penas y alegrías. Gracias en especial a Claudia y Raul por su ayuda con la estadística y R.

Gracias a Serch, por su ayuda con la programación y sobre todo por su amor, su apoyo constante y su comprensión.

Gracias a mis padres, por estar ahí siempre. Le dedico esta tesis a mis abuelos: a mi abuela Teresa que la va a disfrutar aunque no entienda nada y a Antonio, Angelita y Ángel que no han podido verla terminada.

Resumen

La metilación de citosinas es una de las marcas epigenéticas más estudiada. Se sabe que está implicada en la regulación de la expresión génica e interviene directamente en procesos epigenéticos como el imprinting genómico. Por lo tanto, su establecimiento y mantenimiento deben estar finamente regulados en todos los tipos celulares, incluidos aquellos que poseen una dosis extra de ADN. Pese a que se han hecho algunos estudios a este respecto, hasta la fecha todos se han centrado en el efecto de la poliploidía germinal en la regulación epigenética. En esta tesis pretendemos estudiar el efecto que tiene la tetraploidía somática sobre la metilación de citosinas. Para ello empleamos como modelo las neuronas tetraploides corticales recién diferenciadas de ratones de día postnatal 15, siguiendo la línea de investigación del laboratorio. En primer lugar confirmamos que el genoma de estas neuronas se encontraba completamente duplicado. Después analizamos el estado de metilación en sitios CpG encontrando dominios diferencialmente metilados que podrían dar lugar al fenotipo característico de las neuronas tetraploides. El análisis de la metilación de sitios no-CpG reveló mayores niveles de metilación en neuronas tetraploides que podrían indicar que dichas neuronas se encuentran en un estado de maduración más avanzado.

Los genes sujetos a imprinting se encuentran regulados por regiones diferencialmente metiladas (gDMRs) alelo específicas que se establecen en la línea germinal y se cree que se mantienen inalterables en tejidos somáticos. Sin embargo, el análisis de la metilación en estas gDMRs mostró la existencia de hemimetilación hebra-específica en numerosos dominios reguladores de imprinting tanto en neuronas diploides como tetraploides. Además, en el caso concreto del dominio de imprinting de *Peg3*, se describió la presencia de CpGs metil-formiladas en una subpoblación de neuronas tetraploides. Mediante análisis transcripcionales basados en polimorfismos de nucleótido único, vimos que la hemimetilación descrita en el dominio de *Peg3* participa en el mantenimiento de la expresión monoalélica del gen *Peg3*. Nuestros resultados demuestran que las gDMRs pueden ser alteradas en tejidos somáticos y sugieren la existencia de un mecanismo molecular hebra-específico que dé lugar a este tipo de hemimetilaciones.

Abstract

Cytosine methylation is one of the most studied epigenetic marks. It is well established that it has an important role in gene expression regulation and participates in epigenetic mechanisms such as genomic imprinting. Thus, its establishment and maintenance must be tightly regulated in all cell types, including those with an extra dose of DNA. Despite some studies have already been performed on polyploid organisms, they all have been focused on the effect of germinal polyploidy on epigenetic regulation. In this study we have tried to elucidate the effect of somatic tetraploidy in cytosine methylation. To this aim, we used newly differentiated cortical tetraploid neurons from postnatal day 15 mice as a model. We first confirmed that the genome of these neurons is fully duplicated. We then analysed methylation in CpG sites, and found differentially methylated domains that could participate in the establishment of the characteristic phenotype of tetraploid neurons. The analysis in non-CpG sites showed higher levels of methylation in tetraploid neurons which may indicate that they are in an advanced state of maturation.

Imprinted genes are regulated by allele-specific differentially DNA-methylated regions established in the germline (gDMRs), which are supposed to be tightly maintained in somatic tissues. Nonetheless, methylation analysis of these gDMRs showed the existence of strand-specific hemimethylation in several imprinting regulatory domains, both in diploid and tetraploid neurons. Furthermore, in the specific case of the *Peg3* imprinting domain, a small proportion of formylmethylated CpG sites were detected in a subpopulation of tetraploid neurons. Single nucleotide polymorphism (SNP)-based transcriptional analysis indicated that the hemimethylation participate in the maintainance of the monoallelic expression pattern of the *Peg3* imprinted gene. Our results demonstrate that the methylation pattern observed in gDMRs can be modified in somatic tissues and anticipate the existence of a molecular mechanism giving rise to strand-specific hemimethylated CpGs.

Índice

Clave de abreviaturas.....	5
1. Introducción	7
<i>1.1. Epigenética</i>	<i>7</i>
1.1.1. Concepto	7
1.1.2. Metilación del ADN	7
1.1.3. ADN metil-transferasas (DNMTs)	8
1.1.4. Enzimas ten-eleven translocation (TETs).....	9
1.1.5. 5hmC, 5fC y 5caC	10
1.1.6. Reprogramación de la metilación del ADN.....	11
1.1.7. El imprinting genómico	13
1.1.8. La metilación como mecanismo de imprinting genómico.....	14
1.1.9. Técnicas de estudio de la metilación y derivados.....	16
1.1.10. Rasgrf1, Peg3 y Snrpn como modelos de regiones genómicas sometidas a imprinting.....	20
<i>2.1. Tetraploidía</i>	<i>23</i>
2.1.1. Concepto de poliploidía y tipos	23
2.1.2 La tetraploidía somática neuronal.....	23
2.1.3. Tetraploidía e imprinting genómico	24
2. Objetivos.....	27
3. Materiales y métodos.....	29
<i>3.1. Animales</i>	<i>29</i>

3.2. Anticuerpos primarios y secundarios	29
3.3. Extracción de tejido.....	29
3.4. Aislamiento de núcleos.....	29
3.5. Separación de núcleos neuronales	30
3.6. Aislamiento del ADN a partir de núcleos.....	32
3.7. Amplificación del genoma completo e hibridación genómica comparativa mediante chips de ADN.....	32
3.8. Preparación de librerías para WGBS.....	32
3.9. WGBS	32
3.10. Análisis de datos de WGBS	33
3.11. Análisis de los niveles de metilación en las gDMRs de genes sujetos a imprinting	34
3.12. Síntesis de sonda de ADN con 5fC para los ensayos con BH y EA..	34
3.13. Tratamiento del ADN con BH/EA	35
3.14. Conversión con bisulfito.....	35
3.15. NGS del ADN bisulfitado.....	36
3.16. Oligonucleótidos.....	36
3.17. Amplificación, clonación y secuenciación de Sanger del ADN bisulfitado.....	36
3.18. Pirosecuenciación del ADN bisulfitado	38
3.19. Síntesis de ADN no metilado, hemimetilado, metilado y metilformilado para ensayos de corte con HhaI	39
3.20. Digestión de ADN con enzima de restricción sensible a metilación	

<i>seguida de qPCR.....</i>	40
<i>3.21. Hairpin-bisulfite PCR.....</i>	40
<i>3.22. Análisis de la expresión de Peg3</i>	41
<i>3.23. Análisis de datos y estadística</i>	42
4. Resultados	43
<i>4.1. El genoma de las neuronas tetraploides se encuentra completamente duplicado</i>	43
<i>4.2. Mapeo y control de calidad de las lecturas del WGBS</i>	43
<i>4.3. Patrones de metilación global de muestras TETRA y FCX.....</i>	45
<i>4.4. Detección de DMDs.....</i>	50
<i>4.5. Detección de hemimetilación mediante WGBS en gDMRs de genes sujetos a imprinting</i>	51
<i>4.6. Cambios en los niveles de metilación en las gDMRs de Peg3 y Rasgrf1 de neuronas corticales</i>	56
<i>4.7. Hemimetilación de Peg3 en subpoblación de neuronas corticales...</i>	59
<i>4.8. La reducción en la metilación de la Peg3-DMR de neuronas corticales es independiente de la presencia de 5fC.....</i>	63
<i>4.9. Mantenimiento del imprinting del dominio Peg3 en las neuronas corticales.....</i>	66
5. Discusión.....	69
6. Conclusiones.....	77
7. Bibliografía.....	79
8. Anexos.....	101

Clave de abreviaturas

5caC	5-carboxicitosina
5hmC	5-hidroximetilcitosina
5fC	5-formilcitosina
5mC	5-metilcitosina
A	Adenina
AID	Desaminasa Inducida por Activación (del inglés <i>Activation-Induced Deaminase</i>)
VER	Reparación por Excisión de Base (del inglés <i>Base Excision Repair</i>)
BH	Borohidruro sódico
C	Citosina
CAB-Seq	Secuenciación con Bisulfito Asistida Químicamente (inglés <i>Chemical modification-Assisted Bisulphite Sequencing</i>)
CGIs	Islas CpGs (del inglés <i>CpG Islands</i>)
COBRA	Análisis mediante Bisulfito y Restricción Combinadas (del inglés <i>Combined Bisulfite-Restriction Analysis</i>)
DNMTs	ADN Metil-Transferasas (del inglés <i>DNA methyltransferases</i>)
DMDs	Dominios Diferencialmente Metilados en el genoma (del inglés <i>Differentially Methylated Domains</i>)
DMRs	Regiones con Metilación Diferencial en el contexto del imprinting genómico (del inglés <i>Differentially Methylated Regions</i>)
E11	Estadio Embrionario día 11
EDC	1-etil-3-[3-dimetilaminopropil]-carbamida
EA	O-etilhidroxilamina
FACS	Separación de Células Activadas por Fluorescencia (del inglés <i>Fluorescence-Activated Cell Sorting</i>)
fCAB-Seq	Secuenciación con Bisulfito de 5fC Asistida Químicamente (del inglés <i>5fC Chemically Assisted Bisulphite Sequencing</i>)
FISH	Hibridación <i>in situ</i> Fluorescente (del inglés <i>Fluorescent in situ Hybridization</i>)
G	Guanina
gDMRs	DMRs germinales

Clave de abreviaturas

HPLC-UV	Cromatografía Líquida de Alta Eficiencia-Ultravioleta (del inglés <i>High Performance Liquid Chromatography-Ultraviolet</i>)
ICR	Región de Control de Imprinting (del inglés <i>Imprinting Control Region</i>)
IP	Yoduro de Propidio
LC-MS/MS	Cromatografía Líquida acoplada a Espectrometría de Masas en tandem (del inglés <i>Liquid Chromatography coupled with Tandem Mass Spectrometry</i>)
MBDs	Proteínas de Unión a Grupos Metilo (del inglés <i>Methyl-CpG Binding Proteins</i>)
MeDIP-Seq	Secuenciación de ADN Metilado Inmunoprecipitado (del inglés <i>Methylation DNA Immunoprecipitation Sequencing</i>)
NGS	Secuenciación Masiva (del inglés <i>Next Generation Sequencing</i>)
oxBS-seq	Secuenciación con Bisulfito Oxidativa (del inglés <i>Oxidative Bisulphite Sequencing</i>)
P15	Día Postnatal 15
P20	Día Postnatal 20
PGCs	Células del Primordio Germinal (del inglés <i>Primordial Germ Cells</i>)
redBS-Seq	Secuenciación con Bisulfito y Reducción (del inglés <i>Reduced Bisulphite Sequencing</i>)
RRBS	Reduced Representation Bisulfite Sequencing
sDMRs	DMRs somáticas
SNP	Polimorfismo de Nucleótido Único (del inglés <i>Single Nucleotide Polymorphism</i>)
T	Timina
TAB-Seq	Secuenciación con Bisulfito Asistida con TET (del inglés <i>Tet-Assisted Bisulphite Sequencing</i>)
TDG	Timina ADN Glucosilasa (del inglés <i>Thymine DNA Glycosylase</i>)
TETs	Enzimas Ten-Eleven Translocation
U	Uracilo
WAMIDE	atlas web del imprinting genómico y la expresión diferencial en ratón (del inglés <i>web atlas of murine genomic imprinting and differential expression</i>)
X	
WGBS	Secuenciación del Genoma Completo Bisulfitado (del inglés <i>Whole Genome Bisulfite Sequencing</i>)

1. Introducción

1.1. Epigenética

1.1.1. Concepto

Los diferentes tipos celulares que conforman un organismo contienen el mismo genoma y, sin embargo, su morfología celular y sus funciones pueden variar mucho de unos a otros. Las diferentes modificaciones químicas que ocurren en el ADN y en las histonas y que contribuyen a esta diversidad celular se engloban dentro del término epigenética. Estas modificaciones químicas, que no afectan a la secuencia del ADN, tienen además la característica de ser heredables durante la división celular, lo que permite que la información relativa a la expresión génica se transmita a las células hijas (Holliday y Pugh, 1975; Jaenisch y Bird, 2003).

Los mecanismos epigenéticos mejor estudiados incluyen la metilación del ADN y modificaciones post traduccionales de las histonas, como acetilaciones y metilaciones (Bernstein *et al.*, 2007). Algunos autores incluyen a los ARNs no codificantes como mecanismos epigenéticos (revisados en Holoch y Moazed, 2015).

Dado que es la base fundamental de la tesis, en esta introducción nos centraremos en la metilación del ADN y en todo lo que a ella atañe.

1.1.2. Metilación del ADN

Los primeros trabajos que apuntan a la metilación del ADN como un mecanismo epigenético datan de 1975 (Riggs, 1975; Holliday y Pugh, 1975). Los autores de estos trabajos proponían que la metilación del ADN podría tener efectos en la activación e inactivación génica durante el desarrollo y que, además, los patrones de metilación podrían ser heredables.

La metilación del ADN consiste en la adición de un grupo metilo (CH_3), generalmente, a una citosina (C) del ADN, aunque recientemente se han descrito metilaciones en otra base, la desoxiadenosina (Koziol *et al.*, 2016). En los mamíferos, la metilación de la C suele ocurrir en los dinucleótidos CpG y se suele relacionar con el silenciamiento, mediante diferentes mecanismos, de genes que durante el desarrollo participan en la diferenciación celular (Reik *et al.*, 2001). Las islas CpGs (CGIs) son regiones ricas en CpG presentes en

más de la mitad de los genes de vertebrados. Estas CGIs se encuentran principalmente desmetiladas en las células somáticas, pero cuando aparecen metiladas en promotores se relacionan con represión transcripcional (Jones, 2012). En regiones genómicas con un menor número de CpG, la relación entre la metilación del ADN y la transcripción es más compleja y depende en ocasiones del contexto (Jones, 2012).

Aunque las metilaciones de C ocurren en su mayoría en los sitios CpG, también pueden encontrarse en sitios no-CpG (CpA, CpT y CpC), como se ha observado en células madre embrionarias (Ramsahoye *et al.*, 2000; Ziller *et al.*, 2011) y en el cerebro adulto humano y de ratón (Xie *et al.*, 2012; Lister *et al.*, 2013). Se piensa que podría tratarse de una particularidad del tejido nervioso, dado que la metilación en sitios no-CpG es infrecuente o inexistente en otras células somáticas adultas analizadas (Lister *et al.*, 2009; Ziller *et al.*, 2011). Además, la relación inversa entre la metilación en no-CpGs intragénicas y la expresión de mRNA encontrada por Lister y colaboradores (2013) hace pensar que esta modificación podría tener un importante papel en la represión transcripcional.

La metilación del ADN en CpGs es, hasta la fecha, la única marca epigenética cuya heredabilidad durante la mitosis ha sido demostrada (Bird, 2002), pero se sigue investigando sobre la heredabilidad de otros mecanismos epigenéticos.

1.1.3. ADN metil-transferasas (DNMTs)

Las DNMTs son las enzimas encargadas de catalizar la metilación del ADN. Estas enzimas, transfieren un grupo metil de la S-adenosil metionina al carbono 5 de la C, dando lugar a la 5-metilcitosina (5mC). Los patrones de metilación son establecidos *de novo* durante las primeras etapas del desarrollo embrionario y en células del primordio germinal (PGCs), por la DNMT3A y la DNMT3B (Okano *et al.*, 1999). La DNMT3L, que no presenta actividad metiltransferasa por sí misma, actúa como cofactor en esta metilación *de novo* (Bourc'his *et al.*, 2001) (**Figura 1**).

Una vez que los patrones han sido establecidos, la DNMT1, que muestra una mayor afinidad por el ADN hemi-metilado (Hermann *et al.*, 2004; Song *et al.*, 2011), se encarga de mantenerlos después de la replicación del ADN (Leonhardt *et al.*, 1992). En este proceso interviene UHRF1, esencial para el reclutamiento de DNMT1 en la horquilla de replicación (**Figura 1**) (Bostick *et al.*, 2007; Sharif *et al.*, 2007). La DNMT1 es por tanto la metiltransferasa de mantenimiento, pero no es infalible y puede dejar CpGs sin metilar

que, mediante metilación *de novo*, rellenan la DNMT3A y la DNMT3B (Chen *et al.*, 2003).

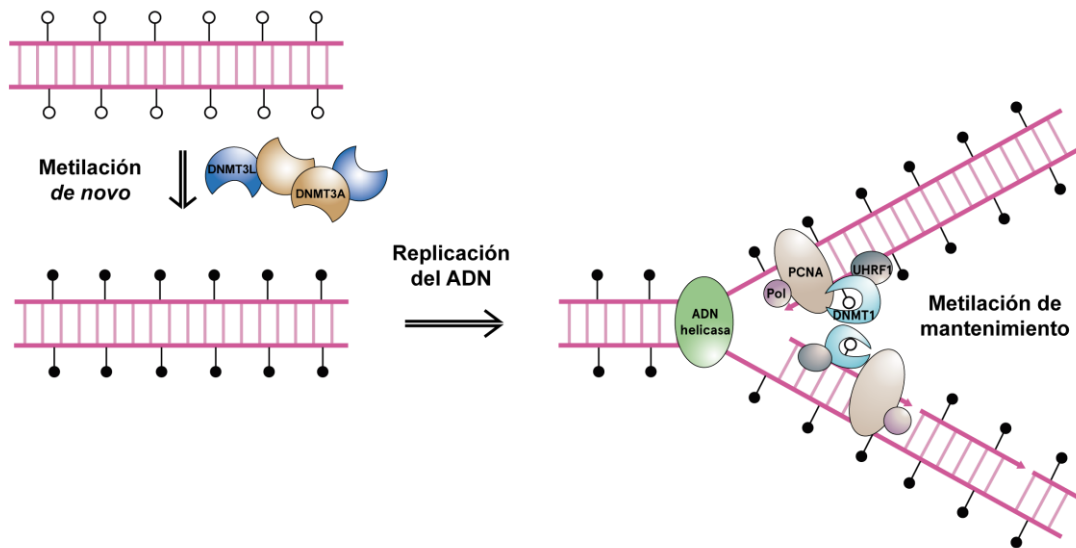


Figura 1. Esquema del mecanismo de actuación de las enzimas DNMT. Las enzimas DNMT3A y DNMT3B, junto con la DNMT3L, que actúa como cofactor, se encargan de establecer la metilación *de novo* en las primeras etapas del desarrollo embrionario y en las PGCs. Durante la replicación del ADN, UHRF1 se encarga de reclutar a la DNMT1 en la maquinaria de replicación para mantener la metilación en ambas hebras del ADN. Basado en Chen y colaboradores (2003).

En ratón, la mayor parte de la DNMT1 se localiza en el núcleo celular. De hecho, DNMT1 contiene numerosas secuencias de localización nuclear dentro del dominio N-terminal (Cardoso y Leonhardt, 1999). Dentro del núcleo celular, la distribución de la DNMT1 depende del ciclo celular (Leonhardt *et al.*, 1992). En la fase G1, se encuentra difusamente distribuida en todo el nucleoplasma. En la fase S temprana, gracias a su dominio de unión a PCNA y al reclutamiento por UHRF1, DNMT1 se dirige a los sitios de replicación. En la fase S-tardía, la DNMT1 se une preferentemente a la heterocromatina pericentromérica en replicación (Schermelleh *et al.*, 2007; Schneider *et al.*, 2013). En la fase G2, la DNMT1 es de nuevo cargada en la heterocromatina pericentromérica por un mecanismo independiente de replicación (Easwaran *et al.*, 2004).

1.1.4. Enzimas ten-eleven translocation (TETs)

Todavía no se ha identificado en mamíferos ninguna desmetilasa capaz de romper directamente la unión covalente entre el grupo metilo y la desoxirribosa de la C, pero estudios recientes investigan la posibilidad de una desmetilación indirecta que implique una oxidación de la 5mC seguida de reparación por excisión de base (BER) (**Figura 2**). La



Hasta la fecha, se han caracterizado, en ratón y en humano, tres miembros dentro de la familia de las proteínas TET: TET1, TET2 y TET3. Los ratones KO para TET3 presentan una mayor incidencia de fallos en el desarrollo que conducen a letalidad (Gu *et al.*, 2011). Sin embargo, los ratones KO para TET1 (Dawlaty *et al.*, 2011) y TET2 (Quivoron *et al.*, 2011) son viables, lo que sugiere que estas proteínas no son esenciales para su desarrollo y que el papel de la TET3 durante la embriogénesis no es compensado por las otras TET. Por otro lado, los dobles mutantes para TET1 y TET2 presentan fenotipos complejos que incluyen defectos en el crecimiento y en el imprinting genómico (Dawlaty *et al.*, 2013).

1.1.5. 5hmC, 5fC y 5caC

10

se vio que la TET1 convertía la 5mC en 5hmC, demostró la existencia de 5hmC en células madre embrionarias de ratón (Tahiliani *et al.*, 2009).

Mientras que la proporción de 5mC es más o menos constante entre los distintos tejidos, la 5hmC varía de unos a otros, siendo más abundante en cerebro (Globisch *et al.*, 2010; Szwagierczak *et al.*, 2010). Además, en células en cultivo, el contenido global de 5hmC disminuye conforme la célula se adapta a las condiciones de cultivo (Nestor *et al.*, 2012). Esta distribución específica de tejido y fases del desarrollo de la 5hmC hacen pensar que la conversión 5mC a 5hmC debe estar finamente regulada.

A pesar de que la función de la 5hmC es todavía desconocida, se cree que puede ser un intermediario en la desmetilación del ADN (Tahiliani *et al.*, 2009; Wu y Zhang 2010). Desmetilación que podría ser pasiva, mediante sucesivas rondas de división celular, ya que la 5hmC apenas es reconocida por la DNMT1 (Valinluck y Sowers 2007), o activa, en la que la 5hmC sería intermediario previo a la BER (Wu y Zhang 2010, He *et al.*, 2011; Maiti y Drohat, 2011) (**Figura 2**).

También se piensa que la 5hmC podría tener cierto papel en la regulación de la transcripción. La acumulación de 5hmC intragénica (Song *et al.*, 2011) podría aumentar la transcripción modificando la accesibilidad de la maquinaria transcripcional o impidiendo la unión de proteínas de unión a grupos metilo (MBDs) represoras (Valinluck *et al.*, 2004; Jin *et al.*, 2010).

Los otros dos derivados de la 5mC, la 5fC y la 5caC, son aún menos abundantes que la 5hmC, pero aun así detectables en células madre embrionarias y distintos tejidos de ratón adulto (Ito *et al.*, 2011). Al igual que en el caso de la 5hmC, una de las funciones propuestas para estos derivados de la 5mC es que funcionen como intermediarios de la desmetilación del ADN (Song y He, 2013, Schiesser *et al.*, 2013) (**Figura 2**). Aunque algunos estudios recientes también señalan que la 5fC es una modificación estable en el ADN que podría tener una función más específica, como señal de reconocimiento de algunas proteínas (Bachman *et al.*, 2015).

1.1.6. Reprogramación de la metilación del ADN

Durante el desarrollo embrionario se producen olas de pérdida y ganancia de metilación del ADN. Después de la fecundación, se produce una desmetilación, mediante los

Introducción

mecanismos pasivos y activos (**Figura 2**) citados en el apartado anterior, de casi todas las marcas epigenéticas del genoma de ambos progenitores (Santos *et al.*, 2002) (**Figura 3**). La eliminación de las marcas epigenéticas heredadas de los gametos lleva al genoma a un estado de totipotencia. Tan solo la metilación de ciertas secuencias, aquellas sometidas a imprinting genómico, así como algunas CGIs (Smallwood *et al.*, 2011), escapan a esta desmetilación como discutiremos en los siguientes apartados.

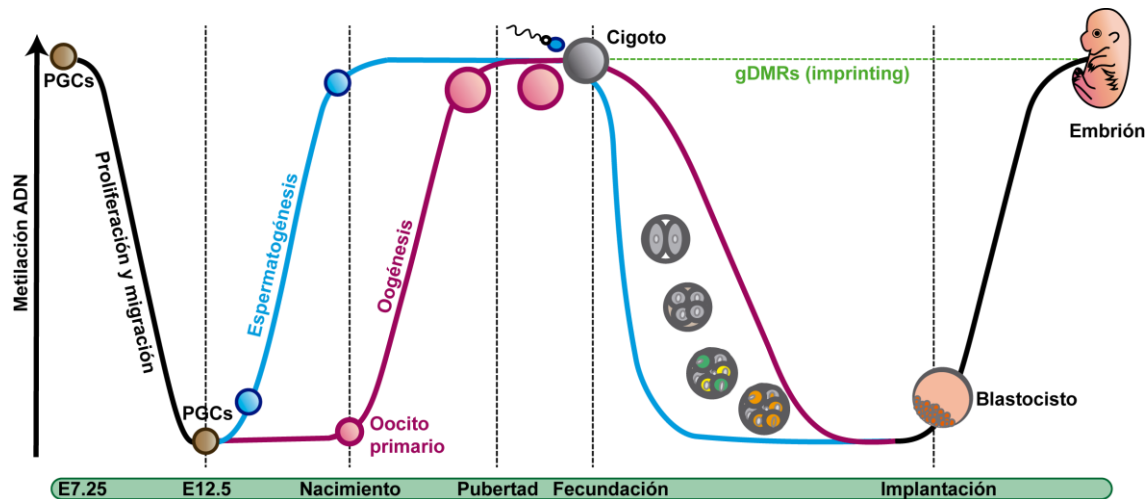


Figura 3. Reprogramación de la metilación durante el desarrollo. Durante la proliferación y migración de las PGCs, se produce un borrado global de la metilación. A continuación, los nuevos patrones de metilación, acordes con el sexo del individuo, se establecen en los gametos. En los machos (línea azul), la metilación *de novo* se produce antes del nacimiento. En hembras (línea granate), por el contrario, la metilación *de novo* se produce después del nacimiento y se completa en la pubertad. Después de la fertilización, se produce un borrado de la metilación del ADN del que escapan las gDMRs de los genes sometidos a imprinting genómico (línea verde discontinua). Mientras que el genoma paterno es borrado rápidamente por un mecanismo de desmetilación activa (línea azul), la desmetilación del genoma materno se produce de manera pasiva (línea granate), diluyéndose en las sucesivas rondas de replicación del ADN. Tras la implantación del blastocisto se establecen los nuevos patrones de metilación en función del tipo celular como parte de la diferenciación. Modificado de Smallwood y Kelsey (2012).

Durante la implantación del blastocito, se produce una metilación *de novo* (Howlett y Reik, 1991, Santos *et al.*, 2002) (**Figura 3**), en la que intervienen la DNMT3A y la DNMT3B mencionadas anteriormente. Se sabe muy poco acerca de por qué algunas secuencias de ADN son metiladas y otras no. Lo más probable es que se trate de un proceso dirigido por lo que muchos estudios actuales se centran en identificar secuencias específicas reconocidas por las DNMTs y sus cofactores.

Las PGCs, que se forman en E7.25, presentan patrones de metilación iguales a los de las células del epiblasto de las cuales proceden (Ginsburg *et al.*, 1990), por lo que deben ser reprogramadas, incluyendo las marcas de imprinting genómico. A pesar de que este

proceso de borrado aún no está bien descrito, se piensa que tiene lugar en el periodo entre E11.5 y E13.5, coincidiendo con el borrado del imprinting en E11.5 (Reik *et al.*, 2001; Hajkova *et al.*, 2002; Seisenberger, 2012) (**Figura 3**), y que es el resultado de la combinación de mecanismos de desmetilación pasivos y activos, dependiendo del contexto (Feng *et al.*, 2010; Saitou *et al.*, 2011; Hackett *et al.*, 2012).

En estudios recientes se han identificado una serie de proteínas que promueven la desmetilación activa de *loci* específicos bajo ciertas condiciones. Una de esas proteínas es la desaminasa inducida por activación (AID), cuya deficiencia afecta a la desmetilación global del ADN (Popp *et al.*, 2010). Sin embargo, el fenotipo epigenético que aparece en esta deficiencia es leve, y sugiere la presencia de mecanismos adicionales de desmetilación que pueden compensar la falta de actividad de AID o actuar en diferentes secuencias diana. La desaminación de la 5mC y la C por las desaminasas AID y APOBEC1 pueden iniciar la vía de BER.

La metilación *de novo* en las células germinales se produce a partir de E12.5 y se mantiene de manera progresiva durante la espermatogénesis de los embriones machos (Davis *et al.*, 2000). En el caso de los gametos femeninos, la metilación se produce postnatalmente en los oocitos en desarrollo (Lucifero *et al.*, 2004) (**Figura 3**). Esta reprogramación de la línea germinal es necesaria para la generación de células germinales funcionales y el fallo en una apropiada reprogramación da lugar a infertilidad o al desarrollo de embriones anormales que mueren durante la gestación.

Las células somáticas, por su parte, van a mantener los patrones de metilación adquiridos tras la implantación del blastocito, aunque en algunos estudios realizados con neuronas post-mitóticas se ha visto que se produce una desmetilación activa que podría tener importancia en la regulación génica (Martinowich *et al.*, 2003; Dong *et al.*, 2007).

1.1.7. El imprinting genómico

En los últimos 20 años, uno de los mecanismos de regulación epigenética más estudiado ha sido el imprinting genómico. Este mecanismo se basa en que la copia materna o paterna de algunos genes se encuentra diferencialmente metilada, dando lugar a una expresión monoalélica del alelo materno o paterno, dependiendo del gen concreto (Li *et al.*, 1993; Ferguson-Smith *et al.*, 1993; Stöger *et al.*, 1993). La pérdida del imprinting después de la

inhibición de la DNMT1 es prueba de que la metilación del ADN es indispensable para este mecanismo de regulación epigenética (Li *et al.*, 1993).

Hoy en día, se han identificado más de 150 genes sujetos a imprinting genómico (Morison *et al.*, 2005), la mayoría de los cuales se encuentran organizados en clústers que se regulan de manera coordinada por una única región de control imprinting (ICR) (Edwards y Ferguson-Smith, 2007; Spahn y Barlow, 2003). Casi todos los clústers contienen al menos un gen no codificante y múltiples genes codificantes que regulan el desarrollo embrionario la placentación y diferentes procesos post-natales, así como el comportamiento (Reik y Walter, 2001).

Los mecanismos a través de los cuales una ICR controla la expresión génica en su clúster respectivo son variados y siguen investigándose, incluyendo los ARN no codificantes y la interacción entre la metilación del ADN, las histonas modificadas y otras proteínas (Barlow y Bartolomei, 2014).

La importancia de una adecuada expresión de los genes sujetos a imprinting para un correcto crecimiento y desarrollo queda patente en embriones que contienen sólo genoma materno o paterno, los cuáles no se desarrollan de manera normal (Solter, 1988) y en enfermedades humanas como el síndrome de Prader-Willi, el síndrome de Angelman o el síndrome de Silver-Russell, ocasionados todos ellos por una pérdida de imprinting (Butler, 2009).

1.1.8. La metilación como mecanismo de imprinting genómico

La metilación alelo específica de las regiones diferencialmente metiladas (DMRs) de las ICRs se ha propuesto como el principal mecanismo de imprinting al regular la expresión monoalélica de muchos genes sujetos a este proceso (Tremblay *et al.*, 1997; Shemer *et al.*, 1997; Lopes *et al.*, 2003; Yoon *et al.*, 2005; Kagami *et al.*, 2010).

En las PGCs, la metilación de las DMRs es establecida por la DNMT3A y DNMT3L (Bourc'his *et al.*, 2001; Hata *et al.*, 2002; Kaneda *et al.*, 2004; Okano *et al.*, 1999), si bien es verdad que la mayoría de las metilaciones ocurren en los oocitos y un menor número en las espermatogonias. Estas DMRs, generadas durante el desarrollo de la línea germinal, se conocen como DMRs primarias o germinales (gDMRs) y su estado de metilación se mantiene tras la fertilización, a pesar de la ola de desmetilación que se produce en ese

momento (Olek y Walter, 1997; Warnecke *et al.*, 1998) (**Figura 3**), y durante toda la vida del individuo (Reik y Walter, 2001). Esto es así salvo excepciones, en las que se producen pérdidas o ganancias de metilación en gDMRs con implicaciones patológicas (revisadas en Feinberg *et al.*, 2002) o como posibles mecanismos de control de la dosis génica expresada (Ferrón *et al.*, 2011).

De manera general, se asume que la adquisición de la metilación en las gDMRs es un proceso activo. Sin embargo, estudios recientes sugieren que también puede existir un mecanismo de protección frente a la metilación y un mantenimiento en regiones metiladas que den lugar a esta metilación diferencial (Kelsey y Feil, 2013).

Además de la acción de la DNMT1, que metila las hebras de ADN recién sintetizadas durante la proliferación, parece ser que el reconocimiento de secuencias por diferentes factores confiere protección a las gDMRs frente a la reprogramación post fertilización, como la ZFP57. Todas las ICRs conocidas presentan dominios de unión a ZFP57, una proteína con dedos de zinc que recluta metiltransferasas (Schultz *et al.*, 2002; Quenneville *et al.*, 2011) que permiten el mantenimiento de la metilación en uno de los alelos. Los ratones deficientes para ZFP57 muestran pérdida de imprinting en muchos *loci* (Li *et al.*, 2008). Lo que protege al alelo no metilado de esta metilación *de novo* no se tiene tan claro, aunque se especula con que en dicho proceso intervengan modificaciones de histona o estructuras en lazos R (Ooi *et al.*, 2007; Ginno *et al.*, 2012).

Por otro lado es fundamental para el desarrollo de la siguiente generación que los patrones de imprinting genómico sean borrados de las PGCs, como se ha explicado anteriormente, y se establezcan las nuevas marcas de imprinting que reflejen el sexo del embrión (**Figura 3**). Además, en las PGCs se reprime el mantenimiento de la metilación que lleva a cabo la DNMT1 tras la replicación del ADN, probablemente al reprimirse UHRF1, la proteína que recluta la DNMT1 en la horquilla de replicación (Kagiwada *et al.*, 2013) (**Figura 1**).

Además de las gDMRs anteriormente descritas, existen DMRs secundarias o somáticas (sDMRs), aquellas que se establecen después de la fertilización y que pueden ser tejido-específicas (Takada *et al.*, 2002; Bhogal *et al.*, 2004). Las sDMRs suelen establecerse en promotores o en regiones intragénicas y su metilación suele corresponder con silenciamiento génico alelo específico (Kobayashi *et al.*, 2012).

1.1.9. Técnicas de estudio de la metilación y derivados

Existen múltiples técnicas para determinar cambios de metilación a nivel global, incluyendo cromatografía líquida de alta eficiencia-ultravioleta (HPLC-UV) (Kuo *et al.*, 1980) y cromatografía líquida acoplada a espectrometría de masas en tandem (LC-MS/MS) (Liu *et al.*, 2009). Sin embargo, a continuación, nos centraremos en las técnicas disponibles para estudiar la metilación a nivel de región o bases simples, ya que algunas de ellas han sido empleadas en esta tesis.

Ninguna técnica de secuenciación tradicional permite distinguir entre las 5mCs y las Cs sin metilar. Por ello, la mayoría de las técnicas empleadas en el estudio de la metilación del ADN están basadas en el tratamiento con bisulfito. Al tratar el ADN con este compuesto, la 5mC no sufre ninguna modificación, mientras que las Cs no metiladas son convertidas en uracilos (U), que tras la amplificación por PCR se convertirán en timinas (T). Comparando la secuencia del ADN convertido con el no convertido, se puede inferir el porcentaje de metilación de una C determinada.

En las primeras aproximaciones, como la llevada a cabo por Frommer y colaboradores (1992), el ADN tratado con bisulfito era amplificado, empleando oligonucleótidos específicos para la región de interés, clonado y secuenciado (**Figura 4**). Este método de análisis de la metilación del ADN ha sido el empleado clásicamente en numerosos estudios, pero, la necesidad de clonar las muestras para secuenciarlas, hace que sea una técnica tediosa. Si en vez de secuenciar las muestras mediante el método tradicional de Sanger se recurre a la pirosecuenciación (Ronaghi *et al.*, 1998) (**Figura 4**), el tiempo y el esfuerzo necesarios para llevar a cabo esta técnica disminuyen dado que no es necesario clonar la muestra después de la amplificación por PCR. Además, la robustez del análisis aumenta.

Otra técnica recurrentemente empleada es el análisis mediante bisulfito y restricción combinadas (COBRA) (Xion y Laird, 1997) (**Figura 4**). Tras tratar el ADN con bisulfito y amplificarlo mediante PCR, la muestra se corta con una enzima cuya diana de restricción debe contener un sitio CpG (por ejemplo, BstUI, CGCG). Dependiendo de si la C de ese sitio se encuentra metilada o no, la diana de la enzima se mantendrá o cambiará, respectivamente, y la enzima podrá realizar o no su corte. Los productos de la digestión se corren en un gel y se realiza una cuantificación de las bandas para determinar el porcentaje

de metilación de la CpG. Se trata de una técnica sencilla y económica, pero sólo aporta información de sitios concretos. Si lo que interesa es ver el estado de metilación de regiones más amplias, es necesario combinar esta técnica con las anteriores.

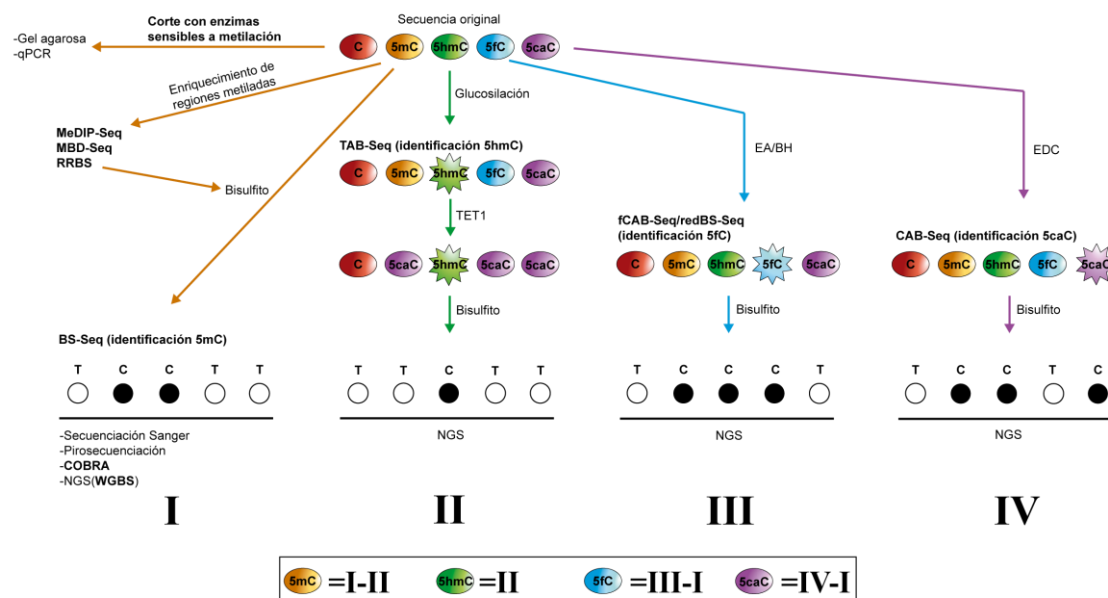


Figura 4. Resumen de métodos para el estudio de la 5mC y sus derivados. La 5mC puede ser detectada por métodos que no requieren el tratamiento con bisulfito, como el corte con endonucleasas sensibles a metilación y su evaluación por gel de agarosa o por qPCR, o con métodos que sí requieren dicho tratamiento. Entre estos últimos, los hay que primero llevan a cabo un enriquecimiento en regiones metiladas con anticuerpos (MeDIP-Seq), proteínas de unión a dominios metilados (MBD-Seq) o tratamiento con enzimas de restricción sensibles a metilación (RRBS); y los hay que directamente tratan el ADN con bisulfito para convertir las Cs no metiladas en Us y estos en Ts. El ADN tratado con bisulfito puede ser cortado con enzimas de restricción cuya diana coincide con una C susceptible de estar metilada (COBRA) o puede secuenciarse, por regiones (secuenciación de Sanger), o por completo, haciendo uso de las nuevas técnicas de NGS (WGBS). Empleando la técnica TAB-seq, la 5hmC se glicosila y queda protegida frente al tratamiento con TET1. Después del tratamiento con bisulfito y la secuenciación, sólo la 5hmC aparecerá como C en la secuenciación. La 5fC puede protegerse empleando EA (fCAB-Seq) o reducirse a 5hmC utilizando BH (redBS-Seq). En ambos casos, el tratamiento con bisulfito no afectará a la nueva forma, que se leerá como C en la secuenciación. Para identificar la 5caC se puede llevar a cabo una aproximación similar, protegiéndola con hidrócloruro de 1-etil-3-[3-dimetilaminopropil]-carbamida (EDC) (CAB-Seq). La 5caC protegida aparecerá como C tras la secuenciación. Realizando las operaciones especificadas en la figura (caja), se obtendrá el porcentaje de cada derivado de la 5mC.

La combinación de la conversión con bisulfito y las técnicas de secuenciación masiva (NGS), ha permitido el análisis de la metilación a lo largo de todo el genoma. Entre las técnicas que incluyen NGS se encuentra la secuenciación del genoma completo bisulfidado (WGBS) (**Figura 4**). De todas las técnicas, esta es la más exacta y reproducible y, además, es la que más información aporta. Sin embargo, se trata de una técnica cara y que entraña

una serie de dificultades a la hora de analizar los datos. La conversión con bisulfito reduce la complejidad del genoma complicando el alineamiento de las lecturas secuenciadas con el genoma de referencia y haciendo necesario el empleo de algoritmos bioinformáticos específicos que permitan dicho alineamiento.

Existen alternativas más económicas a la WGBS como la secuenciación de fracciones del genoma enriquecidas en 5mC. Empleando anticuerpos (Secuenciación de ADN Metilado Inmunoprecipitado, MeDIP-Seq; Zhao *et al.*, 2014), proteínas de unión a dominios metilados (MBD-Seq, Lan *et al.*, 2011) o enzimas de restricción (RRBS, Meissner *et al.*, 2005), se pueden obtener fracciones enriquecidas en 5mC que a continuación son tratadas con bisulfito y secuenciadas utilizando NGS (**Figura 4**). Además de ser más económicas, estas técnicas permiten aumentar la cobertura de la secuenciación (mayor número de lecturas/CpG) y, empleando los anticuerpos adecuados, pueden utilizarse para detectar niveles de 5hmC, indistinguible con el resto de técnicas. Pero tienen menos resolución que el WGBS y además son más susceptibles a las variaciones experimentales, debidas a variaciones en la especificidad de los anticuerpos o a digestiones incompletas del ADN.

Entre las técnicas anteriores, merece especial atención la RRBS. En ésta, el ADN se digiere con la enzima MspI, que reconoce las dianas CCGG y corta tanto el ADN metilado como no metilado, obteniéndose regiones enriquecidas en CpGs. Pero este es precisamente el problema de esta técnica, que se limita a regiones ricas en CpGs dejando muchas zonas donde las CpGs se encuentran más dispersas sin caracterizar. Empleando únicamente MspI, se puede llegar a cubrir hasta el 10% de las CpGs (Boyle *et al.*, 2012). Si además se emplea otra enzima como la ApeKI, la cobertura sube hasta el 20% en células humanas (Wang *et al.*, 2013), pero no llega a cubrir el genoma completo como sí hace el WGBS.

Otra alternativa, más económica que el WGBS, son los microarrays. Empleando microarrays prediseñados se pueden medir los niveles de metilación del ADN en determinadas regiones del genoma. Existen multitud de opciones comerciales disponibles, que ofrecen la posibilidad de estudiar regiones concretas en las que se ha demostrado que la metilación del ADN tiene un efecto funcional, como promotores, elementos reguladores o regiones 3' no traducidas (3'UTRs). No obstante, estos kits suelen estar preparados para estudios con ADN humano, y el diseño de microarrays para ratón puede encarecer los costes además de estar más limitado en el número de CpGs que se pueden detectar.

Pero existen también métodos que no emplean la conversión con bisulfito. Se trata de métodos basados en la digestión selectiva del ADN metilado o no metilado por determinadas endonucleasas (**Figura 4**). Son, en realidad, los primeros métodos empleados para el análisis de la metilación del ADN (Cedar *et al.*, 1979). Mientras la enzima MspI corta en su diana CCGG, indistintamente de si existe metilación o no, la enzima HpaI, que comparte diana con la anterior, corta solamente ADN no metilado. El tratamiento del mismo ADN con ambas enzimas y el análisis del tamaño de los productos de digestión mediante electroforesis es informativo del estado de metilación de CpGs concretas.

Para una cuantificación más fina, se puede digerir el ADN con enzimas de restricción sensibles a metilación (HhaI, McrBC, HpaII, Hin6I y AciI, por ejemplo) y, a continuación, realizar una qPCR con oligonucleótidos específicos que flanqueen el sitio de restricción (Hashimoto *et al.*, 2007) (**Figura 4**). Comparando la amplificación del ADN digerido con una muestra sin digerir, se puede estimar el porcentaje de metilación de una CpG determinada.

Dada la importancia que han cobrado recientemente los intermediarios de la oxidación de la 5mC, se ha hecho necesario el desarrollo de técnicas que permitan su estudio. La detección de estas formas no es una tarea sencilla dada su escasa abundancia (Ito *et al.*, 2011). Para solventar esto, se puede recurrir a técnicas que incluyan un paso previo de enriquecimiento, empleando, como veíamos para 5mC, anticuerpos (DIP-seq, Szulwach *et al.*, 2011; Inoue *et al.*, 2011), marcaje químico selectivo (empleando, por ejemplo, β -glucosiltransferasa, Song *et al.*, 2011), o enzimas de restricción (Szwagierczak *et al.*, 2010).

Existen aproximaciones para el estudio específico de la 5hmC (Secuenciación con Bisulfito oxidativa, oxBS-seq, Booth *et al.*, 2012; Secuenciación con Bisulfito Asistida con TET, TAB-seq, Yu *et al.*, 2012), de la 5caC (Secuenciación con Bisulfito Asistida Químicamente, CAB-seq; Lu *et al.*, 2013) y de la 5fC (Secuenciación con Bisulfito de 5fC Asistida Químicamente, fCAB-seq, Song *et al.*, 2013; Secuenciación con Bisulfito y Reducción, redBS-seq, Booth *et al.*, 2014) a lo largo de todo el genoma (**Figura 4**).

En general, son métodos que se basan en tratamientos químicos del ADN que progen dichas modificaciones del tratamiento posterior con bisulfito. Por ejemplo, en el caso del fCAB-seq, las 5fC se modifican empleando O-etilhidroxilamina (EA), de forma que, al

tratar con bisulfito y secuenciar, se detectarán como 5mC o 5hmC. En el caso de la redBS-seq, las 5fC son reducidas a 5hmC tras tratarlas con borohidruro sódico (BH), que después del tratamiento con bisulfito y la secuenciación se detectarán como C. El problema de estos dos métodos y de los anteriores, es que son necesarias dos secuenciaciones: una en la que se trate el ADN con bisulfito, y otra tratando previamente el ADN con el compuesto químico que proteja/modifique el derivado de la 5mC. El porcentaje de dicho derivado lo dará la diferencia entre ambas secuenciaciones (**Figura 4**). Esto supone un gran gasto económico que no todos los laboratorios pueden permitirse. Además, es necesario emplear controles internos para garantizar la correcta conversión de las muestras, susceptible de variaciones técnicas y biológicas.

1.1.10. *Rasgrfl*, *Peg3* y *Snrpn* como modelos de regiones genómicas sometidas a imprinting

Rasgrfl es un gen sometido a imprinting muy particular. Se expresa solamente a partir del alelo paterno en el cerebro neonatal de ratón hasta el destete (aproximadamente P21). A partir de ese momento, su expresión se vuelve bialélica (Drake *et al.*, 2009). La expresión de *Rasgrfl* está controlada por una gDMR y una serie de repeticiones en el extremo 3' de esta gDMR (40 copias de un elemento formado por 41 nucleótidos) (**Figura 5**). Estas repeticiones dirigen la metilación de la gDMR paterna que regula la transcripción de *Rasgrfl*. La no metilación de la gDMR materna permite que se una CTCF dando lugar a la inhibición en la transcripción del alelo materno (Yoon *et al.*, 2002).

Rasgrfl se expresa principalmente en neuronas del sistema nervioso central (Zippel *et al.*, 2000) con niveles más modestos en otros tejidos somáticos (Font de Mora *et al.*, 2003; Plass *et al.*, 1996). La proteína RasGRF1 es un factor de intercambio de guanina para Ras y Rac (Cen *et al.*, 1993; Innocenti *et al.*, 1999). Activa estas proteínas G en respuesta a corrientes de calcio celular (Farnsworth *et al.*, 1995) o fosforilación de serinas (Mattingly *et al.*, 1999) en vías de señalización de receptores muscarínicos (Mattingly y Macara 1996), disociación de la subunidad heterotrimérica de la proteína G (Kiyono *et al.*, 2000; Shou *et al.*, 1995) y unión de neurotrofinas a receptores TrkA, TrkB y TrkC (MacDonald *et al.*, 1999; Robinson *et al.*, 2005).

RasGRF1 está implicada en múltiples procesos de aprendizaje y en mecanismos de plasticidad, incluyendo la consolidación de memoria dependiente de MAPK y la plasticidad a largo plazo en la amígdala (Brambilla *et al.*, 1997).

El gen de expresión paterna *Peg3* es un gen sujeto a imprinting, descubierto mediante experimentos hechos en ratones partenogénéticos, en los cuáles dicho gen no se expresaba (Kuroiwa *et al.*, 1996). *Peg3* forma parte del clúster que lleva su nombre y que ocupa unas 500kb, en el que se incluyen otros seis genes sujetos a imprinting (**Figura 5**).

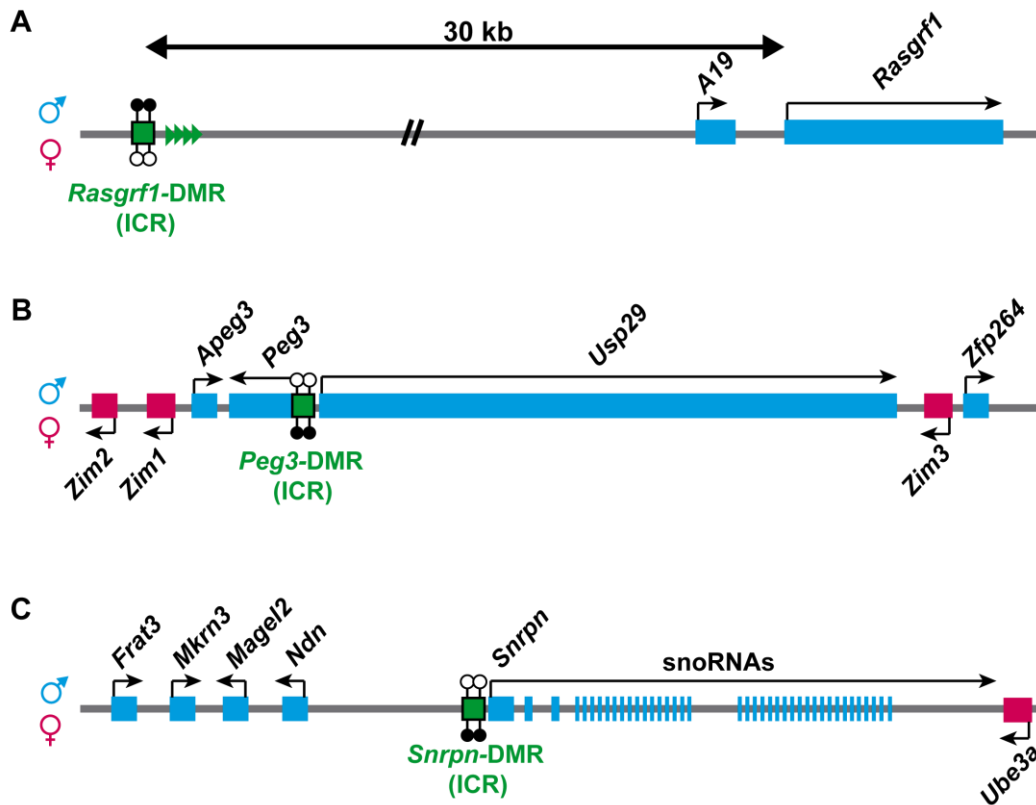


Figura 5. Esquema representativo de la estructura genómica de los dominios de imprinting *Rasgrf1*, *Peg3* y *Snrpn* en ratón. (A) Muestra los genes que constituyen el dominio de imprinting de *Rasgrf1*. La ICR del dominio de imprinting de *Rasgrf1* se encuentra 30kb upstream de dicho dominio. (B) Muestra los genes que constituyen el dominio de imprinting de *Peg3*. (C) Muestra los genes que constituyen el dominio de imprinting de *Snrpn*. Para todos los apartados, los genes expresados a partir del cromosoma paterno se representan en azul, y los expresados a partir del materno en granate. Las gDMRs que forman parte de las ICRs se muestran en verde. Los sitios CpG metilados se representan como círculos negros mientras que los no metilados, se muestran como círculos blancos.

En el dominio de imprinting de *Peg3*, existe una gDMR (Lucifero *et al.*, 2004), y dos sDMRs (Huang y Kim, 2009). De las tres, la gDMR resulta ser la más crítica para la transcripción y el imprinting del dominio *Peg3*. Delecciones de parte de esta gDMR provocan cambios a nivel transcripcional en todos los genes incluidos en este dominio (Kim *et al.*, 2012).

En el cerebro, *Peg3* se expresa principalmente en el hipotálamo (Li *et al.*, 1999), región implicada en el control del comportamiento materno y de la lactancia (Ivanova y Kesley,

2011). Las crías con mutaciones en el locus *Peg3* manifiestan deficiencias en la succión que suelen conllevar letalidad perinatal. Los ratones mutantes adultos de ambos sexos también tienen problemas comportamentales y reproductivos (Kim *et al.*, 2013; Swaney *et al.*, 2008; Swaney *et al.*, 2007). Las hembras con la mutación producen un bajo número de óvulos maduros, dando lugar a camadas reducidas. Además, las hembras preñadas tienen problemas para producir leche y en los cuidados maternos, causando letalidad perinatal entre las crías *wild type*.

En lo que a nivel celular respecta, *Peg3* es un supresor tumoral potencial dada la observación realizada en muchos pacientes de cáncer de ovario y de mama que pierden la expresión de *Peg3* debido a la hipermetilación de su gDMR (Feng *et al.*, 2008). *Peg3* también está implicado en la autofagia de las células endoteliales (Buraschi *et al.*, 2013). Se ha visto que las formas solubles de muchas proteínas estructurales, como la decorina, inducen la autofagia de las células endoteliales con el fin de eliminar aquellas que estén dañadas o sometidas a estrés. *Peg3* parece jugar un importante papel en transmitir esta señal externa a través de VEGFR2 para la transcripción de los dos principales componentes de los autofagosomas: BECLIN1 y LC3. *Peg3* también participa en apoptosis, observándose un incremento del mismo en la apoptosis mediada por p53 en condiciones de hipoxia en el cerebro (Yamaguchi *et al.*, 2002). Como proteína con dedos de cinc, *Peg3* se une a muchos *loci* actuando como represora (Thiaville *et al.*, 2013).

Snrpn forma parte del dominio de imprinting 7C, ortólogo a una región crítica para el Síndrome de Prader-Willi (PWS)/Síndrome de Angelman (AS) en la región 15q11-q13 del cromosoma humano. La región 7C contiene otros genes expresados solo a partir del cromosoma paterno como *Frat3*, *Ndn*, *Magel2*, *Mkrn3* y snoRNAs, y un gen de expresión materna, *Ube3a* (Nicholls y Knepper, 2001). La expresión de los genes contenidos en este dominio se encuentra regulada por una ICR que actúa en cis, situada upstream del gen *Snrpn*. En el locus de *Snrpn* se han localizado dos gDMRs, una en el extremo 5', metilada en el alelo materno (DMR1) y otra en el extremo 3', metilada en el cromosoma paterno (DMR2) (Shemer *et al.*, 1997, Gabriel *et al.*, 1998). La DMR1 es una región de unas 6kb que contiene el extremo 5' y el intrón 1 completo del gen *Snrpn*. La DMR2 es una región de 3.5kb que contiene los exones de 7 a 10 de este gen. Ambas son DMRs primarias cuya herencia es paterno-específica, manteniendo el perfil de metilación desde el gameto (Shemer *et al.*, 1997). Dado que la DMR1 contiene la región de control, el estado de

metilación de la DMR1 no está asociado sólo con la expresión paterno-específica de *Snrpn*, sino que está implicada en el control de imprinting de otros genes de esta región y es en la que nos centraremos en esta tesis.

1.2. *Tetraploidía*

1.2.1. Concepto de poliploidía y tipos

La poliploidía es el aumento de complementos cromosómicos completos de una célula, tejido u organismo. Esta poliploidía puede ser germinal, cuando un fallo en la meiosis de los gametos conduce a que estos presenten un set extra de cromosomas (Comai, 2005) dando lugar a individuos poliploides que pueden ser perfectamente viables, como muchas plantas (Madlung y Wendel, 2013), peces (Leggatt e Iwama, 2003), anfibios (Vernon y Butsch, 1957) e incluso algunos mamíferos (Gallardo *et al.*, 1999).

La poliploidía puede ser también somática, generada mediante diferentes mecanismos revisados en Ullah y colaboradores (2009), y que afecta a poblaciones celulares o tejidos concretos como los hepatocitos (Duncan, 2013), los trofoblastos de placenta (Ullah *et al.*, 2008), los megacariocitos (Vitrat *et al.*, 1998) y los cardiomiocitos (Zanet *et al.*, 2010).

1.2.2. La tetraploidía somática neuronal

Es importante destacar que, debido a la incapacidad de las neuronas para proliferar, no es posible determinar su cariotipo. Por tanto, al indicar que una neurona es tetraploide se hace en sentido laxo, refiriéndose a la cantidad de ADN en vez del número de complementos cromosómicos. El primero en hablar de neuronas tetraploides en vertebrados fue Lowell W. Lapham, quien en 1968 publicó un estudio en el que afirmaba que las células de Purkinje del cerebelo tenían doble cantidad de material genético (Lapham, 1968). Tras él, otros autores se sumaron a la descripción de neuronas tetraploides en vertebrados (Herman y Lapham, 1969; Museridze *et al.*, 1975), pero también surgieron estudios que ponían en duda su existencia (Swartz y Bhatnagar, 1981).

Hoy en día, gracias a los avances en las técnicas de cuantificación del contenido de ADN nuclear se ha podido demostrar la presencia de neuronas tetraploides en distintas zonas del sistema nervioso central de vertebrados (Morillo *et al.*, 2010; López-Sánchez *et al.*, 2011; López-Sánchez y Frade, 2013; López-Sánchez *et al.*, 2017b; Shirazi Fard *et al.*, 2013; Sigl-Glöckner y Brecht, 2017). Utilizando técnicas de citometría estática, citometría de flujo e

hibridación *in situ* fluorescente (FISH) se ha identificado una subpoblación de células ganglionares de retina de pollo que son tetraploides y que inervan la lámina F del estrato gris y fibroso superficial del techo óptico (Morillo *et al.*, 2010). Estas neuronas se generan durante etapas tempranas del desarrollo de la retina y enseguida adquieren marcadores neuronales. Además de en la retina, también se han encontrado neuronas tetraploides en los ganglios de la raíz dorsal, en el cerebelo, en el telencéfalo y en la médula espinal de pollo (López-Sánchez *et al.*, 2011), y en el cerebro del pez cebra (López-Sánchez *et al.*, 2017b).

Mediante estudios con material humano se ha demostrado que el 10% de las neuronas corticales tienen una cantidad de ADN superior a la de las diploides, siendo alrededor de un 2% de ellas tetraploides (Mosch *et al.*, 2007; López-Sánchez *et al.*, 2017). En ratón, se ha visto que existen neuronas tetraploides en la retina (Morillo *et al.*, 2010) y en el cortex (López-Sánchez y Frade, 2013). En el estudio de López-Sánchez y Frade (2013) se vio, además, que las neuronas tetraploides de ratón expresan genes de respuesta al ambiente como *Egr-1* y *fos*, lo que indica que son funcionalmente activas. Asimismo, se observó que un gran porcentaje de las neuronas tetraploides somáticas (60-85%) del córtex cerebral expresan CTIP2, un factor específico de neuronas de proyección de largo alcance (Arlotta *et al.*, 2005).

Por todo esto, y a pesar de que se sabe que la reactivación del ciclo celular en neuronas diferenciadas puede dar lugar a muerte por apoptosis (Vermeulen *et al.*, 2003), la tetraploidía somática neuronal es un proceso regulado que se origina en el desarrollo y que se mantiene a lo largo de la vida adulta en distintas regiones del sistema nervioso central (López-Sánchez *et al.*, 2011; López-Sánchez y Frade, 2013; Morillo *et al.*, 2010).

1.2.3. Tetraploidía e imprinting genómico

Existen diversos estudios realizados con fetos uniparentales (partenogenéticos, androgenéticos y ginogenéticos) en los que se ha demostrado que la expresión de genes sujetos a imprinting se encuentra desregulada (Sotomaru *et al.*, 2001; Sotomaru *et al.*, 2002; Ogawa *et al.*, 2006). En los experimentos realizados por Sotomaru y colaboradores (2002) con fetos androgenéticos, se vio además una reducción parcial en la metilación de la DMR requerida para la expresión monoparental del gen H19.

Por lo tanto, de la misma forma que los genomas uniparentales dan lugar a alteraciones en la metilación de las DMRs y en la expresión misma de genes sujetos a imprinting, cabría

de esperar que la poliploidía, en la cual la dosis genómica se encuentra alterada, diese lugar a alteraciones similares.

En el caso concreto de la tetraploidía, las células presentan dos juegos de cromosomas maternos y dos juegos paternos, a diferencia de lo que ocurre en las células diploides. Una primera aproximación para estudiar la interacción entre la tetraploidía y el imprinting genómico en mamíferos fue la llevada a cabo por Bacquet y colaboradores (2008), que utilizaron como modelo la rata vizcacha roja, que presenta linajes tetraploides (Gallardo *et al.*, 1999). En su estudio, determinaron que tanto la inactivación del cromosoma X como el imprinting genómico se producían de forma similar en células diploides y tetraploides. Sin embargo, en lo que respecta al imprinting genómico, se limitaron a estudiar un solo gen. En particular, vieron que la expresión monoalélica de este se mantenía, pero no hicieron ningún experimento que determinase el estado de metilación de la DMR de dicho gen.

En una investigación más completa y reciente, Yamazaki y colaboradores (2016) estudiaron los niveles de expresión de un total de once genes sujetos a imprinting y el estado de metilación de cuatro DMRs en fetos tetraploides, viendo que, aunque los niveles de expresión de los genes sujetos a imprinting estaban alterados, la expresión monoparental y el estado de metilación de las DMRs se mantenía.

Sin embargo, hasta la fecha no se ha estudiado el estado del imprinting genómico en células tetraploides somáticas.

2. Objetivos

El objetivo global de esta tesis es estudiar si la tetraploidía somática está ligada a cambios epigenéticos, en concreto cambios en los patrones de metilación de las neuronas tetraploides de la corteza cerebral murina y, más específicamente, sobre genes sujetos a imprinting expresados en el sistema nervioso.

Para ello, los objetivos concretos son los siguientes:

- Verificar la duplicación completa del genoma de las neuronas tetraploides.
- Describir el estado de metilación global, tanto de sitios CpG como de sitios no-CpG, en neuronas tetraploides corticales, y compararlo con el del resto de células de corteza cerebral.
- Describir dominios diferencialmente metilados (DMDs) en el genoma de neuronas tetraploides de la corteza cerebral en comparación con el del resto de células de corteza cerebral.
- Determinar si la tetraploidía puede influir en el estado de metilación de las gDMRs que regulan la expresión mono-parental de genes sujetos a imprinting.
- Analizar la presencia de 5fC en dichas gDMRs.
- Determinar si la tetraploidía neuronal puede dar lugar a cambios en la expresión de genes sujetos a imprinting genómico.

3. Materiales y métodos

3.1. Animales

Para los experimentos se emplearon ratones macho C57BL/6 de día postnatal 15 (P15) y estadio embrionario 11 (E11) según la descripción de Kaufman (1992). En embriones, el sexo se determinó mediante PCR estándar con oligonucleótidos específicos del cromosoma Y (ver **Tabla 1**, para consultar la secuencia y la temperatura de *annealing*). Se emplearon también ratones DBA/2J suministrados por The Jackson Laboratory, que fueron cruzados con ratones C57BL/6J para producir ratones P15 con un *background* genético híbrido. Los hígados empleados en esta tesis procedían de ratones macho C57BL/6 de día postnatal 20 (P20). Todos los procedimientos experimentales fueron aprobados por el comité ético del Consejo Superior de Investigaciones Científicas (CSIC), de acuerdo con las leyes de la Unión Europea.

3.2. Anticuerpos primarios y secundarios

Para la separación de núcleos neuronales mediante separación de células activadas por fluorescencia (FACS), se empleó el anticuerpo primario monoclonal de ratón anti-NeuN (Millipore) y el anticuerpo secundario alexa fluor 488 anti ratón hecho en cabra (Abcam) a concentraciones de 1/1000 y 1/500 respectivamente.

3.3. Extracción de tejido

Los ratones P15 se sacrificaron por decapitación y se les extrajo rápidamente el cerebro para colocarlo en PBS. Las cortezas se disecaron bajo la lupa evitando la contaminación de estructuras adyacentes como el hipocampo, el bulbo olfatorio o el estriado, y se congelaron rápidamente en hielo seco para su procesamiento posterior. En el caso de los ratones de estadios embrionarios, se disecó el telencéfalo completo y se congeló en hielo seco. Para los hígados de ratones P20, se disecó uno de los lóbulos y seguidamente se congeló en hielo seco.

3.4. Aislamiento de núcleos

Las cortezas o hígados de ratón se homogeneizaron en 2 ml de tampón de aislamiento de núcleos a 4°C, compuesto por PBS libre de DNasas con 0,1% de Triton X-100 (Sigma) e inhibidor de proteasas (Roche). Los núcleos celulares se aislaron mediante disgregación

mecánica utilizando un homogenizador Dounce. El tejido sin disociar se eliminó mediante centrifugación a 200 g durante 1,5 minutos a 4°C. El sobrenadante se diluyó tres veces con PBS y se centrifugó a 400 g durante 4 minutos a 4°C. Se descartó el sobrenadante con los restos celulares y el *pellet* se incubó a 4°C en 50 µl de PBS durante 30 minutos. Los núcleos se resuspendieron mediante agitación e inversión manual del tubo. La integridad y pureza de los núcleos purificados se analizó con el microscopio tiñendo con 100 ng/ml de DAPI.

3.5. Separación de núcleos neuronales

La inmunotinción de los núcleos neuronales se llevó a cabo añadiendo conjuntamente los anticuerpos primarios y secundarios a 1 mL de suspensión de núcleos aislados sin fijar que contenía 5% FCS y 1,25 mg/ml de BSA. En las muestras control, no se añadió el anticuerpo primario. La mezcla se incubó O/N a 4°C en oscuridad. Antes de la separación por FACS, los agregados nucleares fueron eliminados usando un filtro de 40 µm (Miltenyi Biotec) y se ajustó el volumen a 3 ml con PBS conteniendo yoduro de propidio (IP) (Sigma) y RNasa I libre de DNasas (Sigma) a una concentración final de 40 y 25 ug/mL, respectivamente. La mezcla se incubó 30 minutos a temperatura ambiente. La integridad de los núcleos y la especificidad de la inmunotinción se comprobaron utilizando un microscopio de fluorescencia. Los núcleos aislados del tejido hepáticos fueron filtrados y teñidos con IP como se acaba de describir, pero no se les realizó ninguna inmunotinción.

La separación de núcleos se llevó a cabo con un citómetro FACSAria (BD Biosciences) equipado con un láser doble de argón (488 nm) y helio-neon (633 nm). Los filtros de emisión utilizados fueron el BP 530/30 para el Alexa 488 y el BP 616/23 para el PI. Los datos se analizaron con el software FACSDiva (BD Biosciences). Se aplicó compensación electrónica para corregir el solapamiento entre los espectros de emisión de los distintos fluorocromos tal y como se describe en Herzenberg y colaboradores (2006). Se fijaron las ventanas apropiadas, en primer lugar, para eliminar los restos celulares, fácilmente distinguibles por su incapacidad para incorporar IP, y a continuación para eliminar los agregados, como se describe en Nunez (2001) (**Figura 6**). Después se seleccionó la fracción positiva para NeuN que se representó frente a la intensidad de IP para establecer las ventanas de separación de los núcleos diploides y tetraploides (**Figura 6**).

Mientras que para conseguir poblaciones puras de núcleos neuronales diploides es suficiente con una ronda de separación, para conseguir poblaciones enriquecidas de núcleos tetraploides es necesario llevar a cabo dos rondas de separación. Los núcleos celulares separados en la primera y segunda ronda se recogieron en tubos Protein LoBind (Eppendorf) y DNA LoBind (Eppendorf) respectivamente. Tras la segunda ronda se obtuvo una población con un 90% de núcleos neuronales tetraploides (**Figura 7A**). Los núcleos hepáticos se separaron únicamente en base a su contenido de ADN y fue necesaria una única ronda de separación tanto para los diploides como para los tetraploides, consiguiendo en ambos casos poblaciones puras.

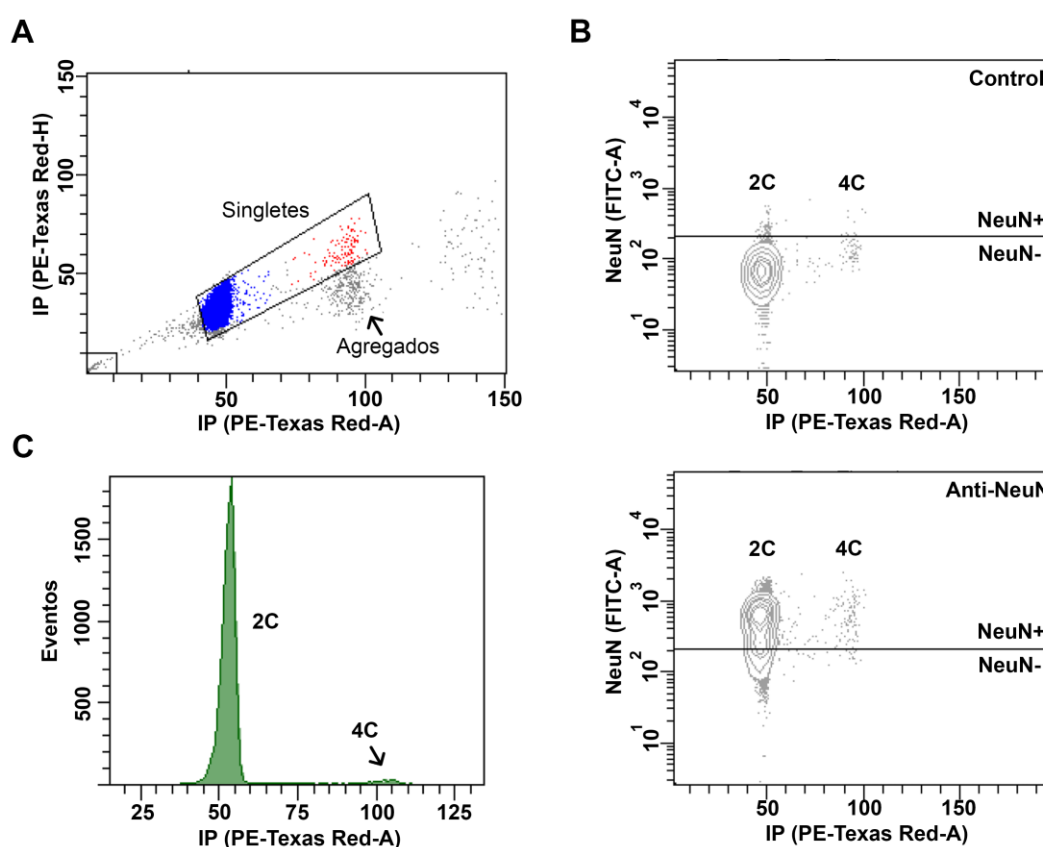


Figura 6. Esquema de ventanas empleado para la purificación de núcleos neuronales diploides (2C) y tetraploides (4C) mediante FACS. (A) Núcleos frescos extraídos de cortex de ratones P15 y marcados con anticuerpo anti-NeuN y con IP se pasaron por el citómetro de flujo. Después de establecer la ventana adecuada en el gráfico que representa el área de la dispersión frontal frente al área de la dispersión lateral (no mostrado), se representó la altura del IP frente al área del IP. En este gráfico, se creó una ventana para separar los núcleos 2C y 4C (singletes) de los dobletes de núcleos diploides (agregados). (B) Se muestran los núcleos 2C y 4C representados en función a sus niveles de marcaje con NeuN (FITC-A) (abajo). Se empleó un control con núcleos a los que no se agregó el anticuerpo secundario (arriba). (C) Histograma con el número de eventos de núcleos NeuN+ 2C y 4C a partir del cuál se hizo la separación. Se emplearon estas mismas ventanas para todos los experimentos detallados en este escrito.

3.6. Aislamiento del ADN a partir de núcleos

El ADN genómico se extrajo de los núcleos separados empleando el kit DNeasy Blood & Tissue Kit (Qiagen) siguiendo el protocolo descrito por el fabricante. Para generar cada ADN genómico se emplearon núcleos neuronales de al menos dos ratones.

3.7. Amplificación del genoma completo e hibridación genómica comparativa mediante chips de ADN

El ADN genómico de núcleos neuronales diploides y tetraploides de corteza de ratones P15 se amplificó utilizando el kit GenomePlex WGA2 (Sigma Aldrich) siguiendo las indicaciones del fabricante. El ADN amplificado con este kit se limpió con fenol cloroformo siguiendo un protocolo estándar. Para la hibridación genómica comparativa mediante chips de ADN se empleó el kit SurePrint G3 Mouse CGH Microarray 180K (Agilent Technologies) con el protocolo facilitado por el fabricante. Este kit contiene sondas que cubren el genoma completo con un espaciado de 10.9 Kb.

3.8. Preparación de librerías para WGBS

Para la preparación de las librerías se empleó el kit Pico Methyl-Seq Library Prep (Zymo Research) siguiendo las instrucciones del fabricante. Las librerías generadas con este kit no son direccionales y tanto la hebra positiva como la negativa se encuentran representadas en ellas. Se prepararon dos librerías a partir de dos ADN genómicos generados independientemente. Se emplearon los index primers compatibles con el sistema TruSeq de Illumina y facilitados en el kit.

3.9. WGBS

La secuenciación del genoma completo bisulfitado se encargó a la unidad de secuenciación del Centro Nacional de Análisis Genómico (CNAG). Allí, en primer lugar, se chequeó la calidad de las librerías empleando el kit DNA 7500 (Agilent Technologies) para el BioAnalyzer 2100 (Agilent Technologies). A continuación, se llevó a cabo la secuenciación *paired end* (2x100pb), en la que se secuencian ambos extremos de los fragmentos de la librería. Para ello se empleó la plataforma Illumina HiSeq 2500. El número de lecturas secuenciadas y la proporción de lecturas alineadas se muestra en la tabla de resultados (**Tabla 2**).

3.10. Análisis de datos de WGBS

El mapeo de las lecturas y el *methylation calling* se llevó a cabo usando el *pipeline* GemBS (<http://statgen.cnag.cat/gemBS/>) desarrollado por el grupo de Heath y colaboradores para el análisis de datos de WGBS. En nuestro caso, se generaron dos genomas de referencia a partir del genoma de ratón *mm10*: C2T, en el que todas las Cs se convirtieron en Ts, y G2A, donde todas las Gs se convirtieron en adeninas (A). El alineador GEM3 (Marco-Sola *et al.*, 2012), incluido en el *pipeline* GemBS, sustituyó las Cs por Ts en la primera lectura de cada par, y las Gs por As en la segunda. Seguidamente se mapearon las lecturas, permitiendo cuatro discrepancias por lectura respecto a la referencia. Se tuvieron en cuenta aquellos pares de lecturas cuyos dos miembros se mapearon con una orientación consistente y no existían con ninguna otra configuración posible.

La estimación del estado de metilación de las Cs se llevó a cabo utilizando el BScall, basado en un modelo Bayesiano e incluido dentro del *pipeline* GemBS. Para cada posición, se calculó el *Phred score*, un valor de calidad de identificación de los nucleótidos secuenciados. Para los análisis posteriores se seleccionaron los sitios CpG donde ambas bases se identificaron como homocigotas con un *Phred score* de al menos 20 (correspondiente a un error en la estimación del genotipo $\leq 1\%$) y donde había más de 5 lecturas.

Las Cs no-CpGs se identificaron como dos nucleótidos adyacentes donde el genotipo del primer nucleótido se identificó como C con una confianza alta y el segundo como no G, también con una confianza alta. Se determinó que estas Cs no-CpGs estaban metiladas si presentaban al menos dos lecturas no convertidas, al menos ocho lecturas informativas del estatus de metilación y una metilación mayor que 0 con $p < 0.05$.

La calidad de los datos obtenidos con WGBS se evaluó empleando la librería RnBeads (Assenov *et al.*, 2014). Dichos datos fueron anotados utilizando los datos de GENECODE para ratón M14 (GRCm38.p5) con software específico diseñado en el CNAG. Para la localización de las CpGs relativa a los genes se emplearon las siguientes categorías: Promotor (1,5kb por encima del extremo 5'), Exon, Intron y regiones intergénicas. Debido a la presencia de sitios de inicio de transcripción alternativos y a regiones que contienen más de un gen, a algunas CpGs se les asigna más de una anotación. Para la localización

relativa a las CGI se emplearon las siguientes categorías: Islas (dentro de la CGI), *Shore* (0–2 kb del borde de las CGI) y *Shelf* (>2 kb–4kb del borde de la CGI).

Para identificar DMDs (denominados así para diferenciarlos de las DMRs de genes sujetos a imprinting) se empleó el paquete DSS, disponible en Bioconductor (Wu *et al.*, 2015). Los parámetros fijados fueron que la DMD contuviese al menos 3 CpGs con un 30% de variación en el porcentaje de metilación, empleando la opción de suavizado de los niveles de metilación de 500pb. DSS detectó las DMDs que mostraban diferencia significativa entre ambos grupos usando un p valor de 0,05 con el test de Wald.

Los DMDs se anotaron como se ha especificado anteriormente. A continuación, se empleó la herramienta de clasificación funcional de genes DAVID (<http://david.abcc.ncifcrf.gov>) que permite asociar funciones biológicas a una lista de genes derivada de otros análisis, en este caso, la detección de DMDs.

3.11. Análisis de los niveles de metilación en las gDMRs de genes sujetos a imprinting

Se recurrió al atlas web del imprinting genómico y la expresión diferencial en ratón (WAMIDEX) para localizar las gDMRs de los genes sujetos a imprinting descritas en la literatura hasta la fecha (Schultz *et al.*, 2008). Se trata de una web en la que figuran las coordenadas de las gDMRs de los genes sujetos a imprinting encontradas en diferentes estudios. Decidimos centrarnos en las gDMRs y utilizar aquellas coordenadas que cubrían una mayor extensión genómica, para tener una visión lo más amplia posible. Una vez localizadas, se empleó la rutina *intersect* del paquete Bedtools para el análisis de los datos obtenidos con WGBS (Quinlan y Hall, 2010). Para determinar si existían diferencias en la metilación de una CpG concreta entre la hebra positiva y la negativa se empleó un algoritmo desarrollado por Raineri y colaboradores (2014), que proporciona la probabilidad de que la metilación en ambas hebras sea diferente. Para este análisis, se filtraron las CpGs y sólo se tuvieron en cuenta aquellas para las que existían lecturas en ambas hebras.

3.12. Síntesis de sonda de ADN con 5fC para los ensayos con BH y EA

Para comprobar la eficacia de los ensayos con BH y EA se sintetizó una sonda de ADN que contenía 5fC tal y como describen Booth y colaboradores (2014). Para ello se

amplificó un molde de ADN (CTCACCCACAACCACAAACAAATTTAATACGATTA AATAATATTAATATATTATCGATTAGTAGGTTAAGTAAGGGTATTTGATGTGAT GGGTGGTATGG) en una reacción de PCR que contenía dATP/dTTP/dGTP (0.2 mM de cada uno, Biotools) junto con 5fCTP (0.2 mM, TriLink), ADN Polimerasa DreamTaq (1.25 unidades, Thermo Fisher Scientific) y los oligonucleótidos 5fC forward y 5fC reverse (0.4 μ M de cada uno, ver **Tabla 1** para secuencia). La PCR se realizó bajo las siguientes condiciones: 3 minutos a 95°C, seguido de un único ciclo de 30 segundos a 95°C, 30 segundos a 57°C, y 15 segundos a 72°C, y una extensión final de 15 minutos a 72°C.

3.13. Tratamiento del ADN con BH/EA

El ADN genómico y la sonda de ADN con 5fC se redujeron usando BH como habían descrito Booth y colaboradores (2014). 5 μ l de una solución 1M de BH (Merck), preparada en el momento, se añadió a 100-200 ng de ADN disuelto en 15 μ l de agua. La reacción se agitó y centrifugó y se mantuvo a temperatura ambiente en oscuridad durante 1 hora con las tapas abiertas. Se agitó y centrifugó cada 15 minutos para eliminar las burbujas generadas. Transcurrida la incubación, la reacción se bloqueó con 10 μ l de una solución que contenía 750 mM de acetato sódico (Sigma) a pH 5.0, incubando 10 minutos a temperatura ambiente o hasta que dejaron de liberarse burbujas.

La protección con EA del ADN genómico y la sonda de ADN con 5fC se llevó a cabo como previamente habían descrito Song y colaboradores (2013), con 10mM de EA (Sigma-Aldrich) en 100 mM de tampón MES, pH 5.0. La reacción se incubó durante 2 horas a 37°C y después se limpió el ADN empleando el kit Genomic DNA Clean & Concentrator™-10 (ZymoResearch). A continuación, las muestras se trataron con bisulfito como se describe más adelante.

3.14. Conversión con bisulfito

Un volumen de 20 μ l de ADN genómico se convirtió con el kit EpiTect Bisulfite (Qiagen) siguiendo las instrucciones del fabricante. Para los experimentos con BH y EA se empleó un protocolo ligeramente modificado por Booth y colaboradores (2014), en el que 100-200 ng de ADN (tratados o sin tratar con BH/EA) disueltos en 30 μ L agua se mezclaron con 80 μ L de mezcla de bisulfito y con 30 μ L de tampón protector del ADN. Esta reacción se sometió a dos rondas de tratamiento con bisulfito y se limpiaron siguiendo el protocolo indicado por el fabricante para muestras fijadas con formalina y embebidas en parafina.

3.15. NGS del ADN bisulfitado

La sonda de ADN con 5fC fue tratada con BH o EA y posteriormente convertida con bisulfito como se describe más arriba. Se prepararon las librerías mediante amplificación con PCR utilizando ADN Polimerasa (Biotools) y los oligonucleótidos 5fC forward post-bisulfite y 5fC reverse post-bisulfite (ver **Tabla 1** para la secuencia). La secuencia del adaptador Illumina adecuado se encontraba incluida en el extremo 5' de los oligos empleados. A continuación, se llevó a cabo la secuenciación *paired end* (2x300pb) de la librería. Para ello se empleó la plataforma Miseq. Las secuencias se alinearon empleando el software MiSeq Reporter (Illumina) y los alineamientos en formato BAM se visualizaron empleando la plataforma Integrative Genomics Viewer.

3.16. Oligonucleotidos

Los oligonucleótidos empleados en los experimentos pueden encontrarse en la **Tabla 1**. Su especificidad fue validada mediante análisis BLAST y por reacciones de PCR.

3.17. Amplificación, clonación y secuenciación de Sanger del ADN bisulfitado

Entre los métodos empleados para confirmar los resultados de WGBS se encuentra el método clásico empleado en muchos estudios de imprinting genómico (Lucifero *et al.*, 2002, Li *et al.*, 2004, entre otros). Este método consiste en el diseño de oligonucleótidos específicos para ADN convertido con bisulfito que flanqueen un número variable de CpGs dentro de la DMR de los genes de interés, en nuestro caso *Snrpn*, *Peg3* y *Rasgrf1* (las secuencias y temperaturas de *annealing* de los primers empleados se encuentran indicadas en la **Tabla 1**). Los oligonucleótidos diseñados dan lugar a amplicones que contienen trece sitios CpG en el caso del dominio de imprinting de *Snrpn* (correspondientes a los sitios CpG numerados del cuatro al dieciséis en Lucifero *et al.* 2002), doce CpGs en el del dominio de imprinting de *Peg3* (correspondientes a los primeros doce sitios CpG de Lucifero *et al.*, 2002) y veintinueve CpGs en el dominio de imprinting de *Rasgrf1* (correspondientes a las CpGs descritas en Li *et al.*, 2004).

Cada reacción de PCR de 50 µl contenía 2 µl de ADN tratado con bisulfito, 1 µl de cada primer 10 µM, 5 µl de dNTPs 2 mM (Biotools), 5 µl de tampón de PCR 10X (Biotools), 2 µl 50 mM MgCl₂ (Biotools) y 2.5 unidades de ADN Taq polimerasa (Biotools). La PCR se

Tabla 1. Listado de oligonucleótidos				
Primer	Secuencia (5' → 3')	Modificación	Aplicación	Temp. de Annealing (°C)
<i>Peg3</i> up-A	TTTTGTAGAGGATTTTGATAAGGAG		Cloning/ Pirosecuenciación	55
<i>Peg3</i> down-A	CAATCTAATACACCCACACTAAACC		Cloning/ Pirosecuenciación	55
<i>Peg3</i> down-A-biot	CAATCTAATACACCCACACTAAACC	5'-biotinilado	Pirosecuenciación	55
<i>Peg3</i> seq-A	ATGTTTATTTTGGGTTGGTGG		Pirosecuenciación	
<i>Peg3</i> up-B	GAGGAGAAGCGGAGAGATGT		Enzyme-qPCR	60
<i>Peg3</i> down-B	CACAGCACTCTACGCACACA		Enzyme-qPCR	60
<i>Peg3</i> probe	AGACTGCCGAGGTCGG	FAM/TAMR A	Enzyme-qPCR	
<i>Peg3</i> up-C	AGAYGTTGGGGAGTTAGGAG		Hairpin-bisulfite PCR	56
<i>Peg3</i> down-C	YAAAAAATATCCACCCTAAACTAAT AAC		Hairpin-bisulfite PCR	56
<i>Peg3</i> hairpin	GCCGAGTCTGACTTTTTTGTCTCAGACT		Horquilla	
<i>Peg3</i> up-D	AATGGCACATGCCTGGAAC	5'-biotinilado	SNP pirosecuenciación	58
<i>Peg3</i> down-D	CGATGAGTGGCCTTGTGTCA		SNP pirosecuenciación	58
<i>Peg3</i> seq-B	CTCCTGTTCACTTCTTTGAGAGAC		SNP pirosecuenciación	N/A
<i>Peg3</i> seq-C	TGAGGGTCTCACTATGTAGGTGT		SNP pirosecuenciación	N/A
<i>Rasgrfl</i> up	TAGAGAGTTTATAAAGTTAG	5'-biotinilado	Cloning/ Pirosecuenciación	55
<i>Rasgrfl</i> down	ACTAAAACAAAAACAACA		Cloning/ Pirosecuenciación	55
<i>Rasgrfl</i> seq	TAATACAACAACAACAATAACAATC		Pirosecuenciación	N/A
<i>Snrpn</i> up-A	TTAGAGGGGATAGAGATTTTGTATT G		Cloning/ Pirosecuenciación	56
<i>Snrpn</i> down-A	CTAAAATCCACAAACCCAATAAC		Cloning	56
<i>Snrpn</i> down-A-biot	CTAAAATCCACAAACCCAATAAC	5'-biotinilado	Pirosecuenciación	56
<i>Snrpn</i> seq	GTATGTGTAGTTATTGTTTGGGA		Pirosecuenciación	56
Chromosome Y-1	GCATTTGCCTGTCAGAGAGAG		Determinación sexo embriones	58
Chromosome Y-2	ACTGCTGCTGCTTTCCAATA		Determinación sexo embriones	58
5fC forward	MTMAMCCAMAACMAMAAAMA		5fC-containing DNA	57
5fC reverse	CCATACCACCCATCACATCA		5fC-containing DNA	57
5fC forward post-bisulfite	ACACTGACGACATGGTTCTACACTC ACTTACAATCACAAACA		NGS	57
5fC reverse post-bisulfite	TACGGTAGCAGAGACTTGGTCTCCA TACCACCCATCACATCA		NGS	57
HhaI forward	CAGAGGACCCTGACAAGGAG		Ensayo de restricción con HhaI	59
HhaI reverse	AGTTCAGATGGTGTTTGGGG		Ensayo de restricción con HhaI	59

realizó bajo las siguientes condiciones: 5 minutos a 94°C, seguido de 35 ciclos de PCR consistentes en 30 segundos a 94°C, 1 minuto a la temperatura de *annealing* correspondiente a cada oligonucleótido (ver **Tabla 1**), y 1 minuto a 72°C, y 10 minutos adicionales a 72°C.

Los productos fueron purificados con fenol-cloroformo y subclonados en el vector pGEM-Teasy (Promega). Con el kit PureYield Plasmid Miniprep (Promega) se purificaron los clones para su posterior secuenciación con el método Sanger. El número de clones analizados varía en función del número de clones que presentaba el inserto adecuado y aquellos que presentaban la secuencia completa.

3.18. Pirosecuenciación del ADN bisulfitado

Otro método empleado para la confirmación de resultados de WGBS fue la pirosecuenciación del ADN bisulfitado. Para ella se diseñaron oligonucleótidos específicos para *Snrpn*, *Peg3* y *Rasgrf1*, uno de los cuales se encontraba biotinilado en su extremo 5' (ver **Tabla 1** para las secuencias y las temperaturas de *annealing* correspondientes). Cada reacción de PCR de 50 µl contenía 2 µl de ADN tratado con bisulfito, 1 µl de cada primer 10 µM, 5 µl de dNTPs 2 mM (Biotools), 5 µl de tampón de PCR 10X (Biotools), 2 µl 50mM MgCl₂ (Biotools) y 2.5 unidades de ADN Taq polimerasa (Biotools). La PCR se realizó bajo las siguientes condiciones: 5 minutos a 94°C, seguido de 50 ciclos de PCR consistentes en 30 segundos a 94°C, 1 minuto a la temperatura de *annealing* correspondiente a cada oligonucleótido (ver **Tabla 1**), y 1 minuto a 72°C, y 20 minutos adicionales a 72°C.

A continuación, los productos de PCR se unieron mediante la biotina del extremo 5' de los oligonucleótidos empleados a beads de sefarosa cubiertas con estreptavidina (GE-Healthcare). Los complejos formados se sometieron a desnaturalización para dejar una única hebra de ADN a la que se unió el oligonucleótido de secuenciación (16 pmol por reacción) (ver la **Tabla 1**). La pirosecuenciación se llevó a cabo en un PSQTM96MA (Biotage) usando 25 µl del ADN amplificado y los reactivos de PyroMark Gold Q96 (Qiagen). Los análisis de los datos de pirosecuenciación se llevaron a cabo utilizando el software Pyro Q-CpG (Biotage).

3.19. Síntesis de ADN no metilado, hemimetilado, metilado y metilformilado para ensayos de corte con HhaI

Para evaluar la capacidad de corte de la enzima HhaI se preparará ADN que contenía dianas de restricción para dicha enzima en distinto estado de metilación y formilación. Los fragmentos de ADN no metilado y metilado se generaron mediante PCR utilizando como molde ADN genómico extraído de tejido cortical. Cada reacción de PCR contenía dATP/dTTP/dGTP (0.2 mM de cada uno, Biotools) junto con CTP (0.2 mM, Biotools) para el ADN no metilado o 5mCTP (0.2 mM, New England Biolabs) para el ADN metilado, además de ADN Polimerasa DreamTaq (1.25 unidades, Thermo Fisher Scientific) y los oligonucleótidos HhaI forward y HhaI reverse (0.4 μ M each, ver **Tabla 1**) flanqueando los sitios de restricción de HhaI de la gDMR del dominio de imprinting de Peg3 (**Figura 18 B**). La PCR se realizó bajo las siguientes condiciones: 3 minutos a 94°C, seguido de 50 ciclos de 30 segundos a 94°C, 1 minuto a 59°C, y 1 minuto a 72°C, y una extensión final de 20 minutos a 72°C. Los productos de PCR se limpiaron empleando el kit DCC Clean & Concentrator-5 (ZymoResearch) siguiendo las instrucciones del fabricante.

Los fragmentos de ADN hemimetilado y metilformilado se generaron también por PCR utilizando como molde el ADN metilado. Cada reacción de PCR contenía dATP/dTTP/dGTP (0.2 mM de cada uno, Biotools) junto con CTP (0.2 mM, Biotools) en el caso del ADN hemimetilado o 5fCTP (0.2 mM, TriLink) para el ADN hemimetilado además de ADN Polimerasa DreamTaq (1.25 unidades, Thermo Fisher Scientific) y los oligonucleótidos HhaI forward y HhaI reverse (0.4 μ M de cada uno, ver **Tabla 1**). La PCR se realizó bajo las siguientes condiciones: 3 minutos a 94°C, seguido de un único ciclo de 30 segundos a 94°C, 1 minuto a 59°C, y 1 minuto a 72°C, y una extensión final de 20 minutos a 72°C. Los productos de PCR se limpiaron empleando el kit DCC Clean & Concentrator-5 (ZymoResearch) siguiendo las instrucciones del fabricante.

Los amplificones no metilados, hemimetilados, metilados y metilformilados (10 μ l de cada uno) se incubaron durante 1 hora a 37°C en presencia o ausencia de 20 unidades de HhaI (Thermo Fisher Scientific). Las reacciones se inactivaron 20 minutos a 80°C y se corrieron en un gel de agarosa.

3.20. Digestión de ADN con enzima de restricción sensible a metilación seguida de qPCR

Para la cuantificación de CpGs concretas en las gDMRs de los dominios de imprinting de Peg3 y Snrpn se purificó ADN de núcleos neuronales separados como se describió anteriormente, y se llevó a cabo un protocolo modificado de Hashimoto y colaboradores (2007). Para cada muestra de ADN, se prepararon en paralelo dos reacciones que contenían 15 ng de ADN en presencia o ausencia de 20 unidades de HhaI (Thermo Fisher Scientific). Todas las reacciones se incubaron durante 5 horas a 37°C en un volumen final de 20 µl. Después de la digestión, las reacciones se inactivaron a 80°C durante 20 minutos.

Se añadieron 1.14 µl de cada reacción a un volumen de reacción de PCR de 20 µl que contenía 10 µl de master mix TaqMan Universal PCR 2X (Thermo Fisher Scientific), oligonucleótidos a una concentración de 900 nM y una sonda fluorescente a una concentración de 250 nM (ver **Tabla 1** para detalles de oligonucleótidos). Las muestras se amplificaron utilizando un termociclador ABI Prism 7000 (Applied Biosystems) bajo las siguientes condiciones: 2 minutos a 50°C, 10 minutos a 95°C, seguido de 40 ciclos de PCR consistentes en 15 segundos a 95°C y 1 minuto a 60°C. Los outlayers se identificaron mediante el test Q de Dixon y se eliminaron.

3.21. Hairpin-bisulfite PCR

Para confirmar la presencia de CpGs hemimetiladas en el la gDMR del dominio de imprinting de Peg3, se adaptó la técnica descrita por Laird y colaboradores (2004). 50 ng de ADN aislado de núcleos neuronales tetraploides separados se digirieron con 2.5 unidades de BbvI (Thermo Fisher Scientific) durante una hora a 37°C. A continuación, se unió una horquilla de ADN al punto de corte de la enzima empleando 10 unidades de T4 ADN ligasa (Thermo Fisher Scientific). El ADN genómico ligado a la horquilla se sometió a conversión con bisulfito usando el kit Epiect Bisulfite (Qiagen). Para impedir que se produjese renaturalización de las dos hebras del ADN propiciada por la horquilla, se introdujeron dos ciclos extra de desnaturalización a los especificados por el fabricante. Las condiciones de incubación finales fueron: 5 minutos a 95 °C, 25 minutos a 60 °C, 5 minutos a 95 °C, 60 minutos a 60 °C ,5 minutos a 95 °C, 60 minutos a 60 °C, 5 minutos a 95 °C, 80 minutos a 60 °C.

El ADN bisulfitado se amplificó por PCR con los oligonucleótidos correspondientes (ver **Tabla 1** para detalles sobre la secuencia). Cada reacción de PCR de 50 µl contenía 2 µl de ADN tratado con bisulfito, 1 µl de cada primer 10 µM, 5 µl de dNTPs 2 mM (Biotools), 5 µl de tampón de PCR 10X (Biotools), 5 µl 50 mM MgCl₂ (Biotools) y 2.5 unidades de ADN Taq polimerasa (Biotools). La PCR se realizó bajo las siguientes condiciones: 1 minuto a 94°C, seguido de 35 ciclos de 30 segundos a 94°C, 1 minuto a 56°C, y 1 minuto a 72°C, y una extensión final de 20 minutos a 72°C.

Los amplificones fueron purificados con fenol-cloroformo y subclonados en el vector pGEM-Teasy (Promega). Se purificaron los clones con el kit PureYield Plasmid Miniprep (Promega) para su posterior secuenciación con el método Sanger. El número de clones analizados varía en función del número de clones que presentaba el inserto adecuado y aquellos que presentaban la secuencia completa.

3.22. Análisis de la expresión de Peg3

Para determinar si la expresión del gen *Peg3* seguía siendo monoalélica, se localizó un polimorfismo de nucleótido único (SNP) utilizando el *Mouse Genome Informatics tool* (The Jackson Laboratory) (http://www.informatics.jax.org/strains_SNPs.shtml). Este SNP (NCBI accession number rs31442449) se caracteriza por la sustitución de una T (en los ratones C57BL/6J) por una C (en los ratones DBA/2J), en la posición 6712228 del cromosoma 7, dentro del gen *Peg3* (NCBI accession number NC_000073). El ARN total se extrajo utilizando el kit RNAqueous-4PCR (Ambion) a partir de núcleos neuronales diploides y tetraploides de cortex de machos P15 obtenidos en cruzamientos recíprocos entre ratones DBA/2J y C57BL/6J. Los cortex se trataron con Allprotect (Qiagen) antes del aislamiento de núcleos siguiendo el protocolo especificado por el fabricante. Las muestras de ARN se trataron con DNasa I (Thermo Fisher Scientific) para eliminar la posible contaminación de ADN genómico que pudiese existir. La síntesis del cADN se llevó a cabo empleando la Transcriptasa Reversa SuperScript III (Invitrogen) siguiendo las instrucciones del fabricante. Para las PCRs se emplearon oligonucleótidos específicos que flanqueasen el SNP, uno de los cuales estaba biotinilado en su extremo 5' (ver **Tabla 1**). Los productos de se pirosecuenciaron utilizando un oligonucleótido interno de secuenciación adecuado (ver **Tabla 1**) como se indicó anteriormente.

3.23. *Análisis de datos y estadística*

Los cambios en los patrones de metilación de las secuencias obtenidas mediante secuenciación de Sanger se analizaron empleando la herramienta para la cuantificación de la metilación (Quma) descrita por Kumaki y colaboradores (2008) (<http://quma.cdb.riken.jp/>). Este software genera un panel con círculos negros y blancos indicando las CpGs metiladas y no metiladas respectivamente. Los datos cuantitativos se obtuvieron del análisis de ADN por triplicado de al menos tres ADNs diferentes (obtenidos a partir de al menos dos ratones) por cada punto experimental, excepto para los experimentos de BH con hígados de ratones P20, para los que sólo se empleó un ADN por cada punto experimental.

Los análisis estadísticos y los gráficos se llevaron a cabo utilizando el software Graphpad Prism 5 o el paquete estadístico R y librerías como *ggplot2*. Los análisis de metilación global se llevaron a cabo juntando los datos de todos los sitios CpG. Los datos de qPCR se analizaron empleando la versión 1.2.3 del software 7000 System Sequence Detection (Applied Biosystems). El porcentaje de metilación en el ensayo de digestión de ADN con enzima de restricción sensible a metilación seguida de qPCR se estimó dividiendo los valores obtenidos en las alícuotas de ADN digeridas frente a aquellas en las que no se incorporó la enzima. Los valores se obtuvieron de trece experimentos independientes en el caso de diploides y seis en el caso de tetraploides, con sus correspondientes triplicados. Las diferencias significativas se determinaron empleando el test de la t de Student a no ser que se especifique otra cosa.

4. Resultados

4.1. El genoma de las neuronas tetraploides se encuentra completamente duplicado

Antes de realizar cualquier otro análisis, era muy importante demostrar que el genoma de las neuronas tetraploides estaba completamente duplicado. De otro modo, cualquier alteración en el porcentaje de metilación en una región concreta podría ser explicada por la presencia de un número indeterminado de duplicaciones de regiones del ADN. Para ello se purificaron núcleos neuronales diploides y tetraploides de ratones P15 C57BL/6J macho mediante FACS (22000-27000 núcleos aproximadamente, **Figura 7 A**) de los cuales se extrajo el ADN para realizar un ensayo de hibridación genómica mediante chip de ADN. Debido a que la cantidad de ADN obtenida tras el FACS era inferior a la requerida para este ensayo, se llevó a cabo un paso intermedio de amplificación del genoma completo.

Con el ensayo de hibridación genómica mediante chip de ADN se demostró que el genoma de las neuronas tetraploides se encuentra completamente duplicado, de forma que las variaciones en el estado de metilación descritas posteriormente no se deben a duplicaciones o deleciones de cromosomas o de regiones cromosómicas (**Figura 7 B**).

4.2. Mapeo y control de calidad de las lecturas del WGBS

Para el análisis del metiloma de las neuronas tetraploides de corteza de ratones P15 se llevó a cabo WGBS. Se secuenciaron dos réplicas biológicas (AC2503 y AC2504) y sus lecturas se alinearon con el genoma de referencia empleando el *pipeline* GemBS tal y como se detalla en *Materiales y métodos*. En la secuenciación de la muestra AC2503 se obtuvieron 568.694.102 lecturas de las cuales un 32,51% fueron mapeadas. En la secuenciación de la muestra AC2504 se obtuvieron 687.612.628 lecturas de las cuales un 31,95% fueron mapeadas (**Tabla 2**). Los datos obtenidos con las dos replicas fueron combinados una vez que se comprobó la reproducibilidad de las mismas y se procedió a su análisis conjunto. A esta combinación de datos se la denominará de ahora en adelante como TETRA.

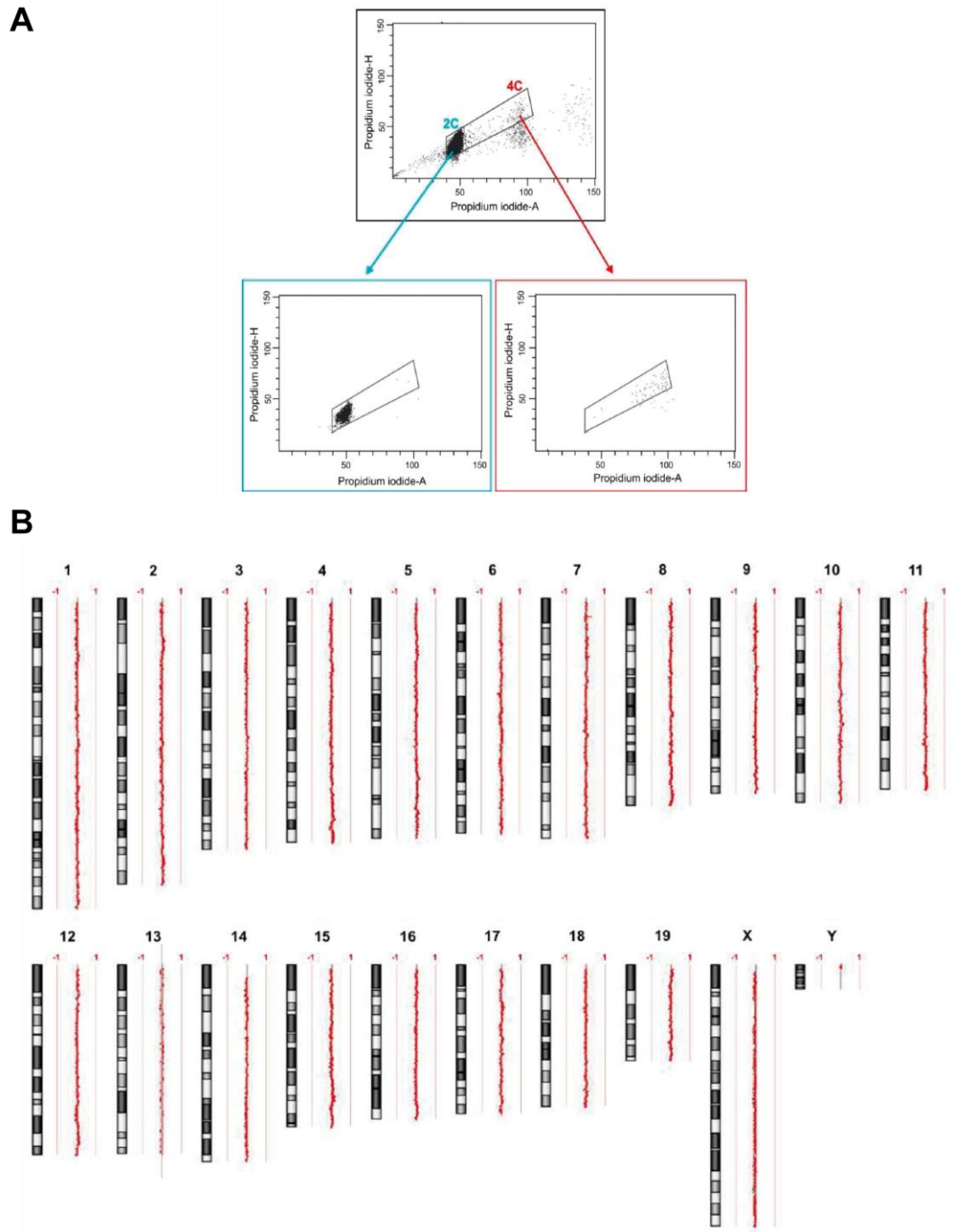


Figura 7. (A) Resultado del FACS para la purificación de núcleos neuronales diploides (2C) y tetraploides (4C). En la gráfica superior se muestran los núcleos frescos extraídos de corteza de P15 marcados con anticuerpo anti-NeuN y con IP (área del IP vs altura del IP). Los núcleos neuronales incluidos en la ventana son separados (como se indica en la **Figura 6**) dando lugar a dos poblaciones neuronales enriquecidas en núcleos neuronales 2C (gráfico de abajo a la izquierda) y 4C (gráfico de abajo a la derecha) **(B) Análisis de hibridación genómica mediante chip de ADN comparando el ADN genómico aislado de núcleos neuronales 2C vs. 4C.** No se observó desviación del ratio respecto a 0 en ninguno de los dominios cromosómicos analizados (trazos rojos), lo que indicaba que el genoma de las neuronas 4C estaba completamente duplicado.

Empleando el mismo *pipeline*, se analizó en paralelo un set de datos obtenido de Lister y colaboradores (2013), correspondientes a corteza frontal (FCX) de ratones P15 machos. Dado el reducido porcentaje de neuronas tetraploides en la corteza (2%, López-Sánchez *et al.*, 2017), los datos de esta muestra pueden ser asimilados a neuronas diploides y células de glía. En esta muestra se obtuvieron 886.855.061 lecturas de las cuales se mapearon un 77,33% (**Tabla 2**).

Tabla 2. Resumen de datos obtenidos en el WGBS						
Nombre	Muestra	Lecturas crudas	Tasa mapeo (%)	#CpGs	#noCpGs	Procedencia
TETRA	AC2503	568.694.102	32,51%	16.264.840	945.441	WGBS
	AC2504	687.612.628	31,95%			WGBS
FCX	FCX	886.855.061	77,33%	16.540.898	1.073.452	Lister <i>et al.</i> , 2013

Las lecturas cubrieron todos los cromosomas por completo (**Figura 8 A**) y por cada sitio CpG se obtuvo una cobertura promedio de 22,55 lecturas en el caso de FCX y de 17,92 lecturas en el caso de TETRA (**Figura 8 B**).

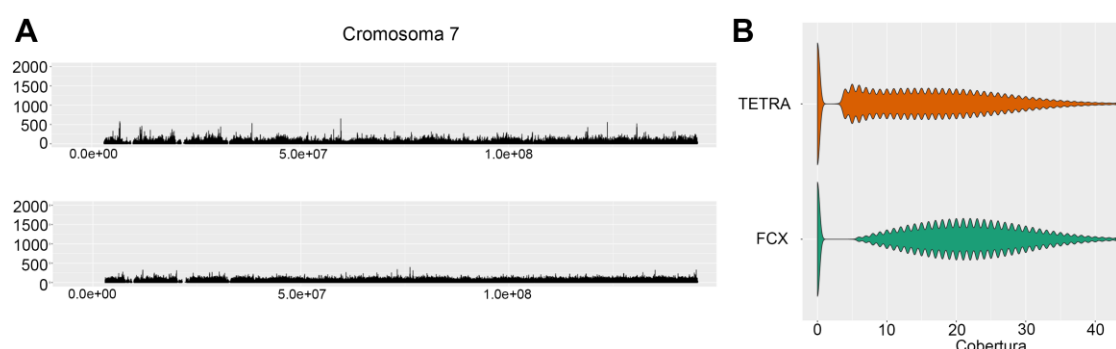


Figura 8. Control de calidad de la cobertura de WGBS de las muestras FCX y TETRA. (A) Figura representativa de la cobertura efectiva de las lecturas en el cromosoma 7 de TETRA (arriba) y FCX (abajo) (B) *Violin plots* que muestran la cobertura de secuenciación para la muestra de TETRA (naranja) y FCX (verde).

4.3. Patrones de metilación global de muestras TETRA y FCX

En primer lugar, se examinó la distribución de la metilación global de las CpGs en ambas muestras, FCX y TETRA (**Figura 9 A**), comprobándose que la mayoría de las CpGs se encontraban hipermetiladas (M, > 70% de metilación) en ambos casos (media de metilación de 81,52% para FCX y 79,06% para TETRA). La cantidad de CpGs parcialmente metiladas (PM, 30-70% de metilación) fue similar en ambas muestras, aunque ligeramente mayor en la de FCX (**Figura 9 A**) mientras que la cantidad de CpGs

Resultados

hipometiladas (NM, <30% de metilación) era mayor en la muestra TETRA que en la muestra FCX.

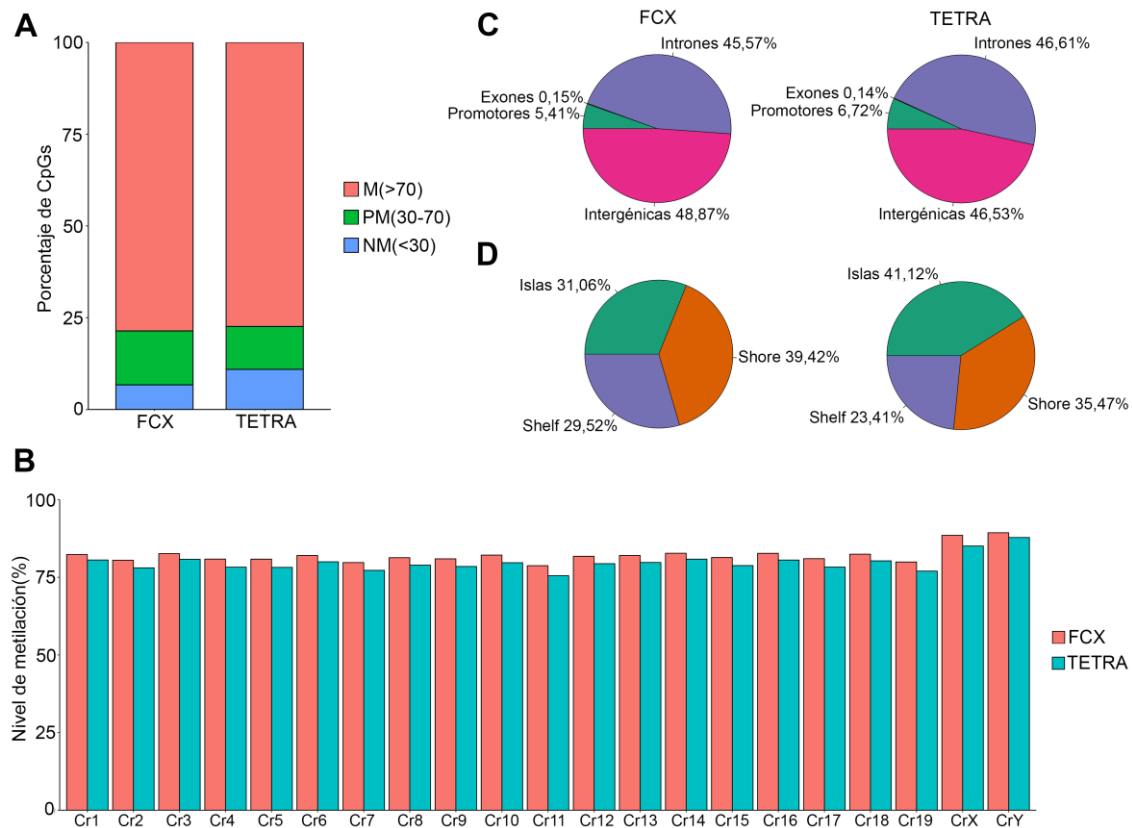


Figura 9. Análisis global de metilación en los sitios CpG de las muestras analizadas mediante WGBS. (A) Estado de metilación global de las muestras de cortex prefrontal (FCX) y neuronas tetraploides separadas por FACS (TETRA) analizadas por WGBS. Se muestra el porcentaje de CpGs metiladas (M, rojo), parcialmente metiladas (PM, verde) y no metiladas (NM, azul). **(B)** Distribución de los niveles de metilación en los cromosomas 1-19, X e Y para ambas muestras, FCX (rojo) y TETRA (turquesa). **(C)** Distribución de sitios CpGs en función de la región funcional (Promotores, Exones, Intrones y regiones intergénicas) para la muestra FCX (izquierda) y TETRA (derecha). **(D)** Distribución de sitios CpGs en CGI (Islas) y regiones adyacentes (*Shore* y *Shelf*) para la muestra FCX (izquierda) y TETRA (derecha).

A continuación, se analizaron los niveles de metilación de los sitios CpGs a lo largo de los cromosomas y se representaron en un histograma para ambas muestras (**Figura 9 B**). Los resultados seguían la tendencia de los niveles de metilación global, con una ligera reducción en los niveles de metilación en la muestra TETRA.

También se analizaron los niveles de metilación de los sitios CpG en las diferentes regiones genómicas. En primer lugar, se realizó un recuento de las CpGs cubiertas con WGBS anotadas para las distintas regiones funcionales y para CGI y regiones adyacentes (*Shelf* y *Shore*). Mientras que el resultado para las regiones funcionales fue muy semejante

en ambas muestras (**Figura 9 C**), en el caso de las CGIs, el número de sitios CpG registrados en la muestra TETRA fue ligeramente mayor que en FCX (**Figura 9 D**).

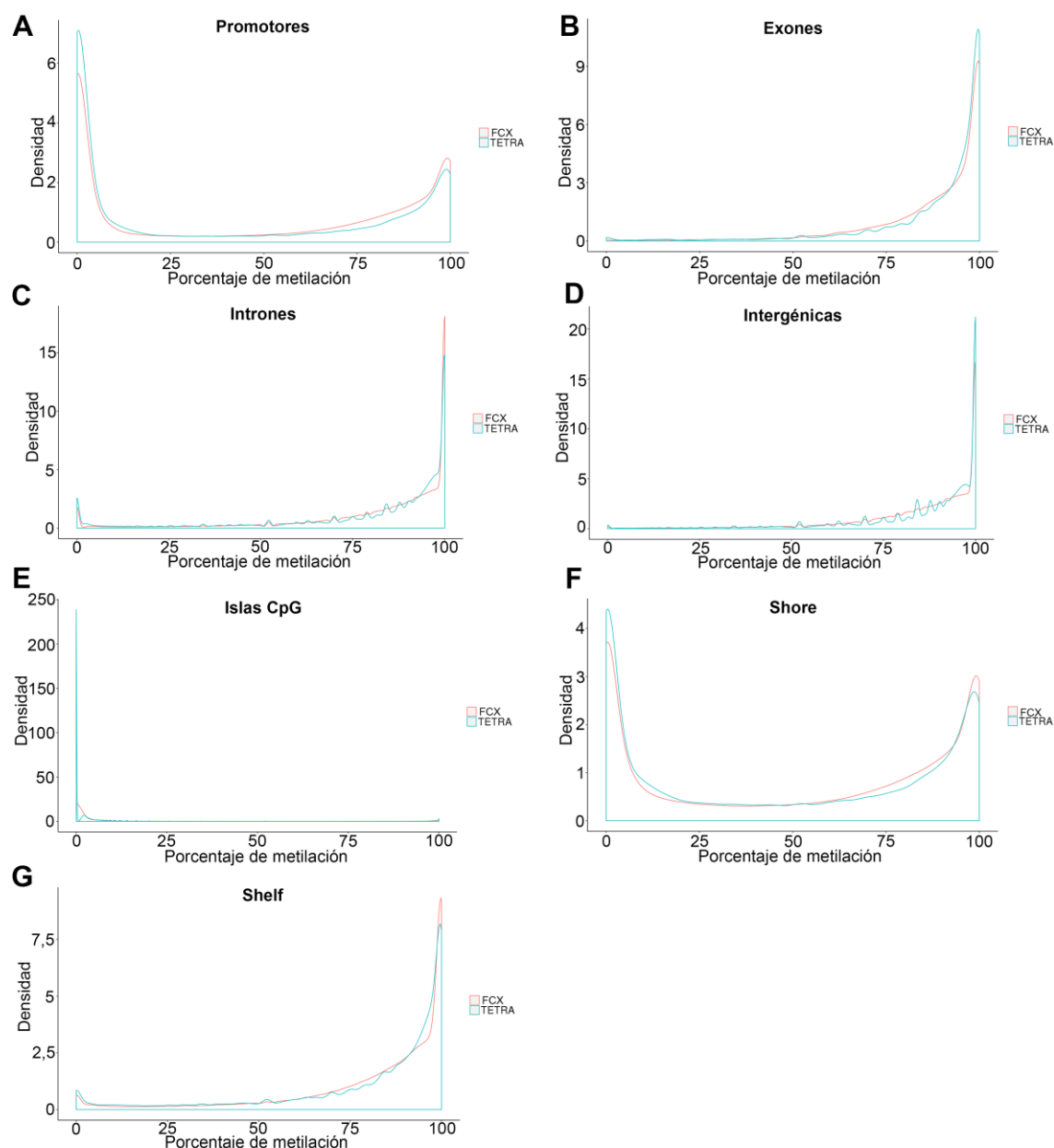


Figura 10. Distribución de niveles de metilación en sitios CpG, en función de diferentes regiones funcionales para FCX (rojo) y TETRA (azul). En las gráficas se aprecia que la distribución de los niveles de metilación era equivalente en (A) promotores, (B) exones, (C) intrones y (D) regiones intergénicas. También se observó la misma distribución en islas CpG (E), *Shore* (F) y *Shelf* (G). Empleando además el test de Kolmogorov-Smirnov no observamos diferencias significativas en ninguna de las regiones analizadas.

A continuación, se analizó la distribución de los niveles de metilación en función de las diferentes regiones genómicas (**Figura 10**) y se comprobó, utilizando el test de

Resultados

Kolmogorov-Smirnov, que no había diferencias significativas en la distribución de los niveles de metilación de ninguna de ellas.

Dado que la metilación en sitios no-CpG es un fenómeno frecuente en neuronas (Lister *et al.*, 2013), decidimos estudiarlo en las neuronas tetraploides. Se secuenciaron aproximadamente 1.073.000 y 945.000 sitios no-CpG en las muestras de FCX y TETRA, respectivamente (**Tabla 2**). La distribución de la metilación estaba desplazada hacia el 0, y la mayoría de las no-CpGs se encontraban hipometiladas en ambos casos (media de metilación de 30,8% para FCX y 31,76% para TETRA, **Figura 11 A**). El contexto más frecuente es CpApC, tanto en TETRA como en FCX, coincidiendo con lo observado en estudios anteriores (Lister *et al.*, 2013) (**Figura 11 C**).

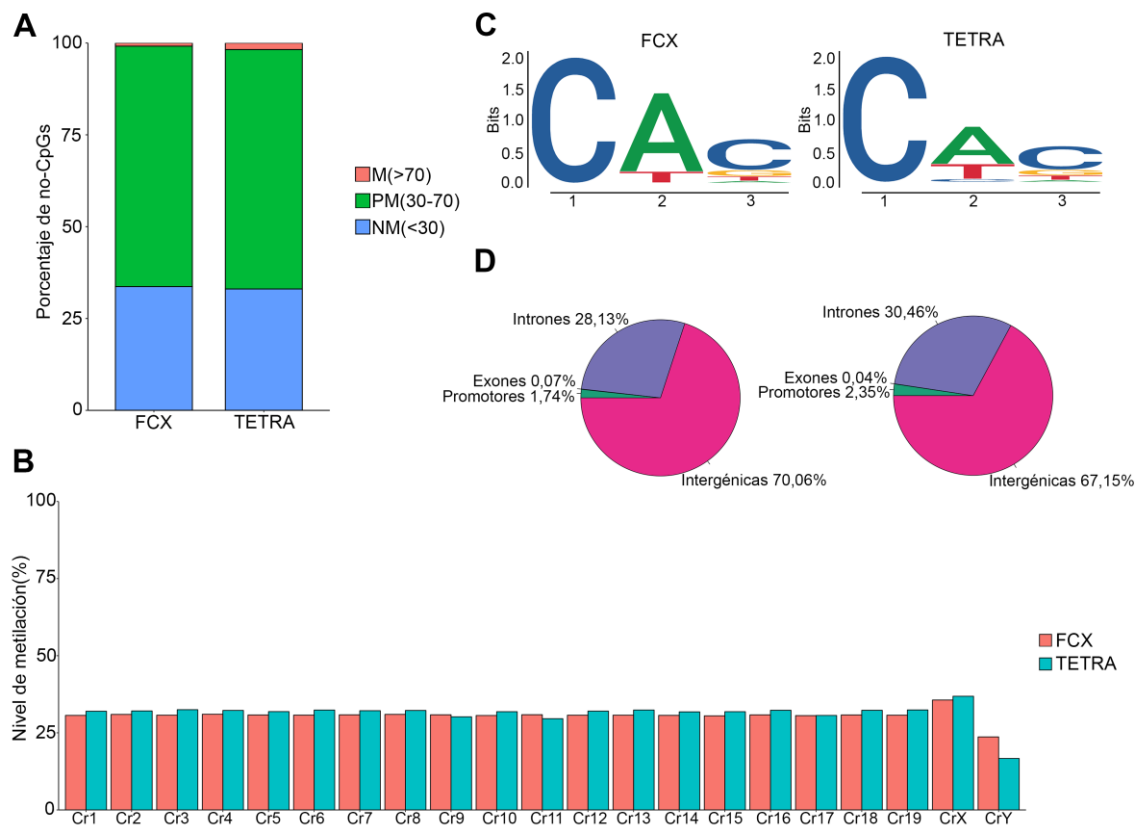


Figura 11. Análisis global de metilación en los sitios no-CpG de las muestras analizadas mediante WGBS. (A) Estado de metilación global de las muestras de corteza prefrontal (FCX) y neuronas tetraploides separadas por FACS (TETRA) analizadas por WGBS. Se muestra el porcentaje de no-CpGs metiladas (M, rojo), parcialmente metiladas (PM, verde) y no metiladas (NM, azul). (B) Distribución de los niveles de metilación en los cromosomas 1-19, X e Y para ambas muestras, FCX (rojo) y TETRA (turquesa). (C) Logo plots del contexto más frecuente para las no-CpGs. (D) Distribución de sitios no-CpG en función de la región funcional (Promotores, Intrones, Exones, y regiones intergénicas) para la muestra FCX (izquierda) y TETRA (derecha).

Como en el caso de los sitios CpGs, se analizaron los niveles de metilación de los sitios no-CpG a lo largo de los cromosomas y se representaron en un histograma para ambas muestras (**Figura 11 B**). En este caso llamó la atención el hecho de que la muestra TETRA mostrara un ligero aumento en los niveles de metilación respecto a la muestra FCX, excepto para el cromosoma Y. En este cromosoma, los niveles de metilación en ambas muestras eran menores a los del resto de cromosomas, y notablemente menores en el caso específico de la muestra TETRA.

Se analizó también la distribución de los sitios no-CpG en las diferentes regiones genómicas (**Figura 11 D**). En este caso, observamos un ligero enriquecimiento de no-CpGs metiladas en regiones no codificantes.

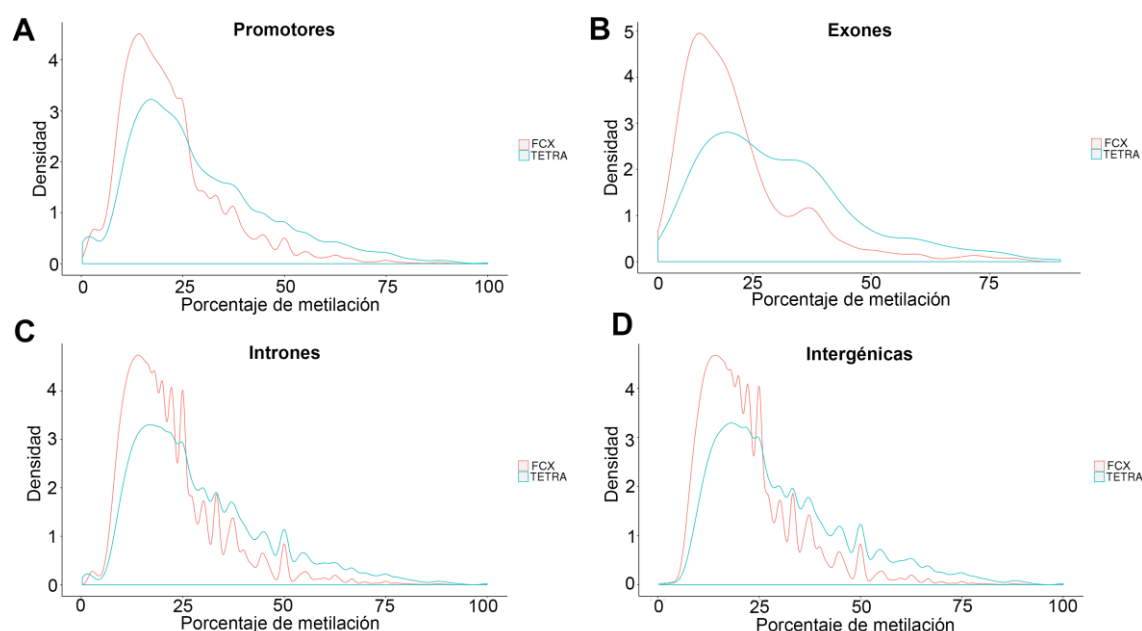


Figura 12. Distribución de niveles de metilación en sitios no CpG, en función de diferentes regiones funcionales para FCX (rojo) y TETRA (azul). En las gráficas se aprecia que la distribución de los niveles de metilación está ligeramente desplazada hacia una mayor metilación en la muestra TETRA en (A) promotores, (B) exones, (C) intrones y (D) regiones intergénicas. Empleando el test de Kolmogorov-Smirnov comprobamos que las distribuciones eran diferentes significativamente (promotores: p-valor < 2.2e-16; exones: p-valor = 8.571e-14; intrones: p-valor < 2.2e-16; regiones intergénicas: p-valor < 2.2e-16).

Después se analizaron los niveles de metilación de las no-CpGs en función de las diferentes regiones genómicas (**Figura 12**), para lo cual se tomaron los datos de los sitios no-CpGs comunes en ambas muestras. En todos los casos, se observó que los niveles de metilación estaban fuertemente desplazados hacia la hipometilación (**Figura 12**), como habíamos visto al analizar la metilación global (**Figura 11 A**). Se comprobó también,

Resultados

utilizando el test de Kolmogorov-Smirnov, que las distribuciones de los valores de metilación eran diferentes significativamente en todas las regiones analizadas (**Figura 12**). La distribución de los valores de metilación estaba desplazada hacia la derecha en el caso de la muestra TETRA (**Figura 12**) denotando mayores niveles de metilación en no-CpGs en neuronas tetraploides tal y como se había observado anteriormente (**Figura 11**).

4.4. Detección de DMDs

Para identificar DMDs entre las muestras FCX y TETRA, empleamos la librería DSS (Wu *et al.*, 2015) con los parámetros especificados en *Materiales y métodos*, y con los datos de metilación de CpGs.

Se detectaron 8.341 DMDs, localizadas mayoritariamente en intrones y seguidamente en promotores (**Figura 13 A**). De las DMDs detectadas con la librería DSS, 4.065 mostraban mayor porcentaje de metilación en la muestra FCX (metilación FCX > metilación TETRA) y 4.375 mostraban mayor porcentaje de metilación en la muestra TETRA (metilación TETRA > metilación en FCX).

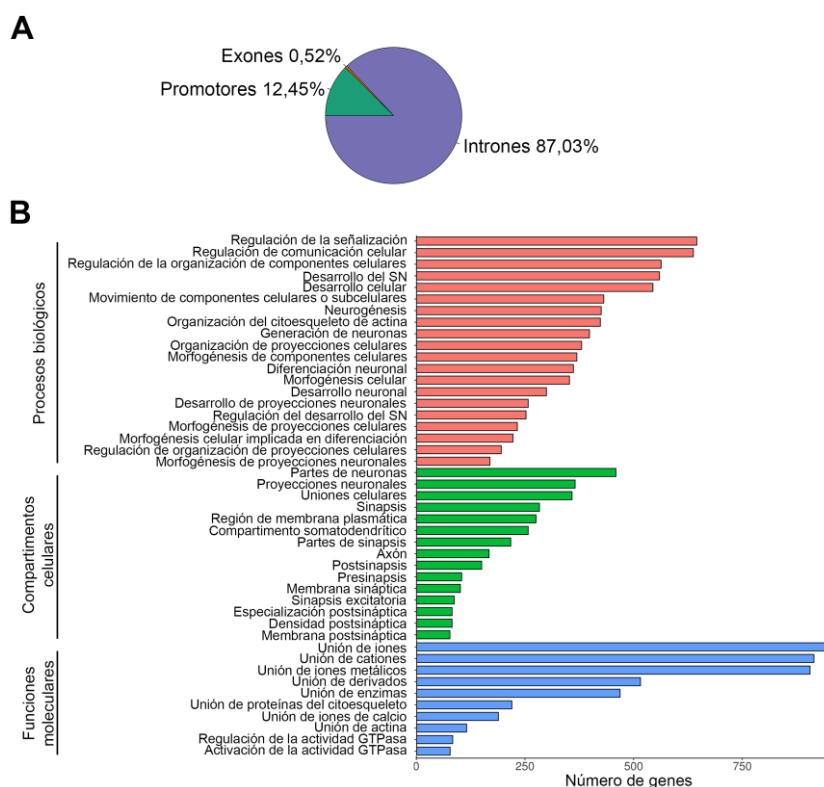


Figura 13. Detección de DMDs entre las muestras FCX y TETRA. (A) Distribución de las DMDs en función de la región funcional (Promotores, Intrones, Exones). La mayoría de las DMDs se encontraron en intrones. (B) Categorías de GO (*Gene Ontology*) enriquecidas en genes con DMDs.

Para intentar encontrar una potencial función biológica a estas DMDs, llevamos a cabo un análisis de enriquecimiento GO (*Gene Ontology*) utilizando la herramienta para anotación funcional DAVID. Todas las DMDs se anotaron en función a tres categorías: procesos biológicos, compartimentos celulares y funciones moleculares. Se seleccionaron las categorías con un valor en el test de Benjamini menor de 0,001 y se representaron los resultados (**Figura 13 B**). Dentro de los procesos biológicos, las DMDs estaban enriquecidas dentro de *regulación de la señalización* (645; 12,67%), *regulación de la comunicación celular* (637; 12,51%) y *desarrollo del sistema nervioso* (559; 10,98%). Dentro de la categoría compartimentos celulares, las DMDs estaban enriquecidas en *partes de neuronas* (459; 9,02%), *proyecciones neuronales* (365; 7,17%) y *uniones celulares* (358; 7,03%). En la categoría de funciones moleculares, las DMDs estaban enriquecidas en *unión de iones* (953; 18,72%), *unión de cationes* (914; 17,96%) y *unión de iones metálicos* (905; 17,77%).

Con los datos obtenidos, podemos decir que las DMDs detectadas podrían participar en el establecimiento del fenotipo característico de las neuronas tetraploides.

4.5. Detección de hemimetilación mediante WGBS en gDMRs de genes sujetos a imprinting

Empleando la base de datos WAMIDEX se localizaron las coordenadas de las gDMRs de los genes sujetos a imprinting descritas en la literatura hasta la fecha (Schultz *et al.*, 2008) y se analizó el estado de metilación de las mismas a partir de los datos obtenidos mediante WGBS. En algunos casos, un clúster de genes sujetos a imprinting puede contener más de una DMR, como es el caso del clúster de Gnas, que incluye las gDMRs Gnas-exon1A, Nesp y Nespa. Lo mismo ocurre en el caso de Gpr1-Zdbf2, que contiene dos gDMRs. En este estudio se analizó el estado de metilación de todas ellas.

Pese a que la cobertura de las gDMRs fue variable entre las muestras FCX y TETRA, de manera general se puede decir que los niveles de metilación en la mayoría de las DMRs se aproximaban al 50% de metilación esperado, tanto en FCX (**Figura 14**) como en TETRA (**Figura 15**). Hay solamente una excepción notable, la de la gDMR de Dlk1-Gtl2, en la que se observaron niveles de metilación altos en todas las CpGs secuenciadas, cercanos al 100%.

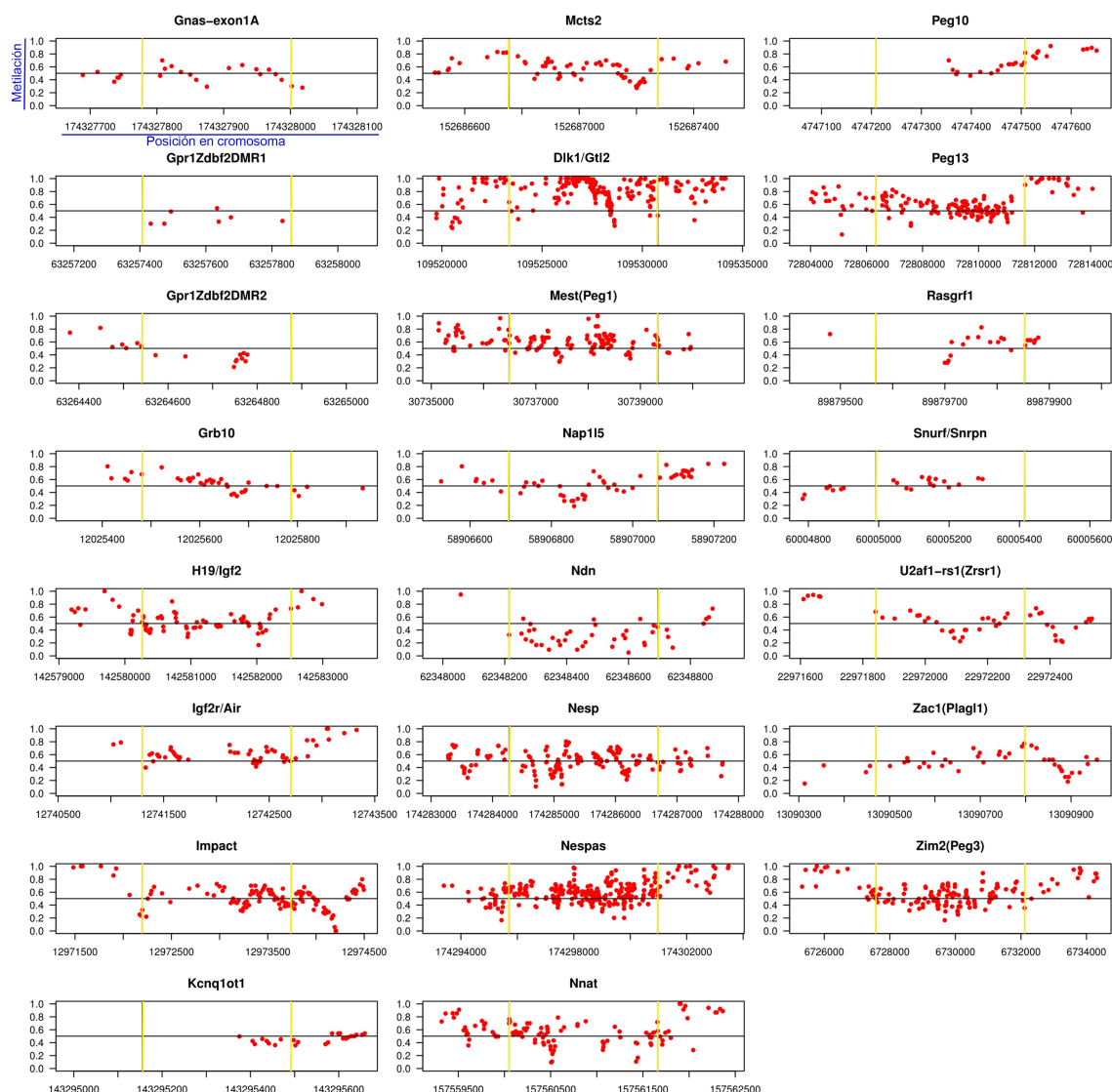


Figura 14. Niveles de metilación global de las gDMRs de genes sujetos a imprinting en FCX. En la imagen se muestran los niveles de metilación de las CpGs analizadas mediante WGBS. Cada punto rojo corresponde con una CpG con su valor de metilación específico. Las líneas verticales amarillas delimitan la gDMR de cada gen concreto. La nube de puntos se ajusta al 50% de metilación en la mayoría de los casos.

A continuación, decidimos analizar los datos de metilación para cada hebra del ADN y determinar si el porcentaje de metilación variaba de una a otra empleando el algoritmo de Raineri y colaboradores (2014) tal y como se describe en *Materiales y métodos*. Con esta aproximación pudimos identificar CpGs dentro de las gDMRs, tanto en FCX como en TETRA, en las que el porcentaje de metilación era diferente en una hebra respecto de la otra en base a los resultados del algoritmo anteriormente citado (Raineri *et al.*, 2014) (Figura 16 y Figura 17).

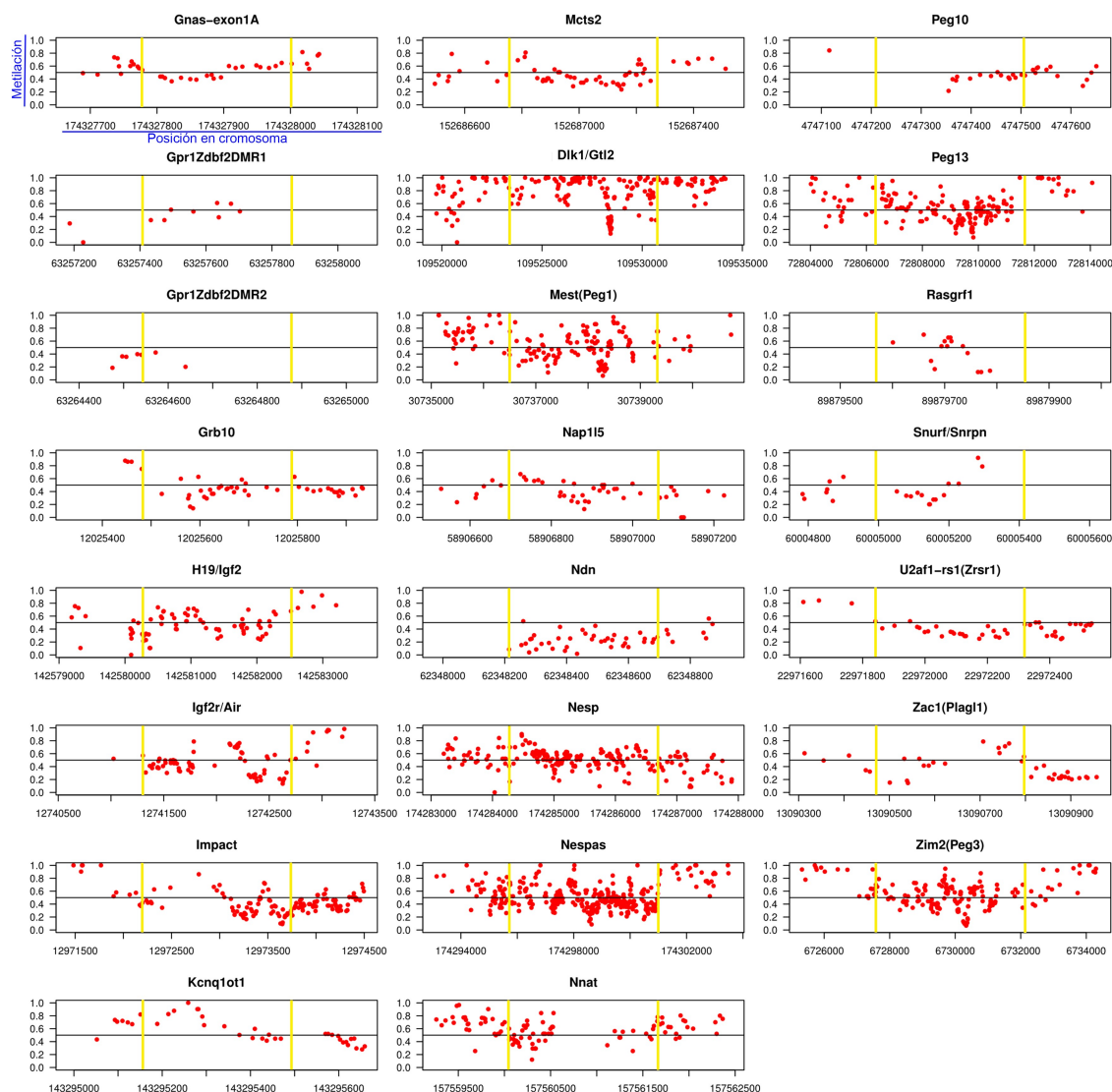


Figura 15. Niveles de metilación global de las gDMRs de genes sujetos a imprinting en TETRA. En la imagen se muestran los niveles de metilación de las CpGs analizadas mediante WGBS. Cada punto rojo corresponde con una CpG y con su valor de metilación específico. Las líneas verticales delimitan la DMR de cada gen concreto. La nube de puntos se ajusta al 50% de metilación en la mayoría de los casos.

Estos resultados muestran que, aunque los valores globales de metilación no parecen estar alterados en las gDMRs, existe hemimetilación en ciertas CpGs de algunos genes sujetos a imprinting, tanto en células de FCX como en neuronas tetraploides.

Resultados





Figura 17. Metilación hebra-específica en las gDMRs de TETRA. Se representan las gDMRs de los genes sujetos a imprinting. Para cada CpG registrada mediante WGBS, al porcentaje de metilación de la hebra positiva se le restó el porcentaje de metilación de la hebra negativa (Dif. Metilación) y se representó el resultado como un punto naranja cuando no existía diferencia significativa entre ambas hebras, o como una cruz morada cuando existía diferencia significativa. Las regiones sombreadas en verde corresponden a las regiones analizadas en experimentos posteriores.

El ADN extraído de estos núcleos se trató con bisulfito y la metilación en una de las dos hebras (determinanda por los oligonucleótidos de elección, ver *Discusión*) se analizó mediante secuenciación tradicional de Sanger y mediante pirosecuenciación, empleando los oligonucleótidos especificados en la **Tabla 1**. Para la pirosecuenciación se emplearon oligonucleótidos internos específicos (**Figura 18**).

En el caso del gen *Rasgrf1*, tanto la secuenciación de Sanger (**Figura 19 A**) como la pirosecuenciación (**Figura 20 A y B**) mostraron que el porcentaje de metilación era significativamente mayor al 50% esperado, tanto para las células diploides como para las tetraploides, confirmando lo observado con los resultados de metilación hebra-específica obtenidos mediante el WGBS (**Figura 16 y Figura 17**). También se pirosecuenció como control el ADN genómico de E11 convertido con bisulfito, ya que en ese estadio, el telencéfalo contiene principalmente precursores neurales, y se vio que en esa etapa la metilación estaba en torno al 50% (**Figura 20 A y B**).

En *Peg3* se observó una disminución significativa en los niveles de metilación respecto al nivel teórico esperado, tanto por secuenciación de Sanger (**Figura 19 B**) como por pirosecuenciación (**Figura 20 C y D**). Se comprobó además que los niveles de metilación también eran significativamente menores que los del neuroepitelio telencefálico de E11. Cabe notar además que, para este gen, la disminución fue significativamente mayor en el caso de las células tetraploides.

Para comprobar que esta disminución observada en los niveles de metilación de la *Peg3*-DMR de las neuronas tetraploides no se debía a su condición de poliploides se empleó ADN de hepatocitos de ratones P20. Los hepatocitos tetraploides y diploides se separaron mediante FACS con un protocolo muy parecido al empleado para los núcleos neuronales. En estos, se observó una ligera disminución en los niveles globales de metilación de la *Peg3*-DMR de tetraploides respecto a los de diploides (**Figura 20 E**) pero, al analizar los niveles de cada CpG de manera individual, no se observó ninguna diferencia significativa (**Figura 20 F**).

Para *Snrpn* no se observaron diferencias significativas respecto al 50% esperado ni con la secuenciación de Sanger (**Figura 19 C**) ni con la pirosecuenciación (**Figura 20 G y H**) lo cual también estaba en consonancia con lo encontrado en el WGBS.

Resultados

Los datos aquí presentados validan los ya obtenidos con el WGBS y muestran que, al menos en neuronas corticales de ratones P15, existen variaciones en los niveles de metilación de las gDMRs de algunos genes sujetos a imprinting.

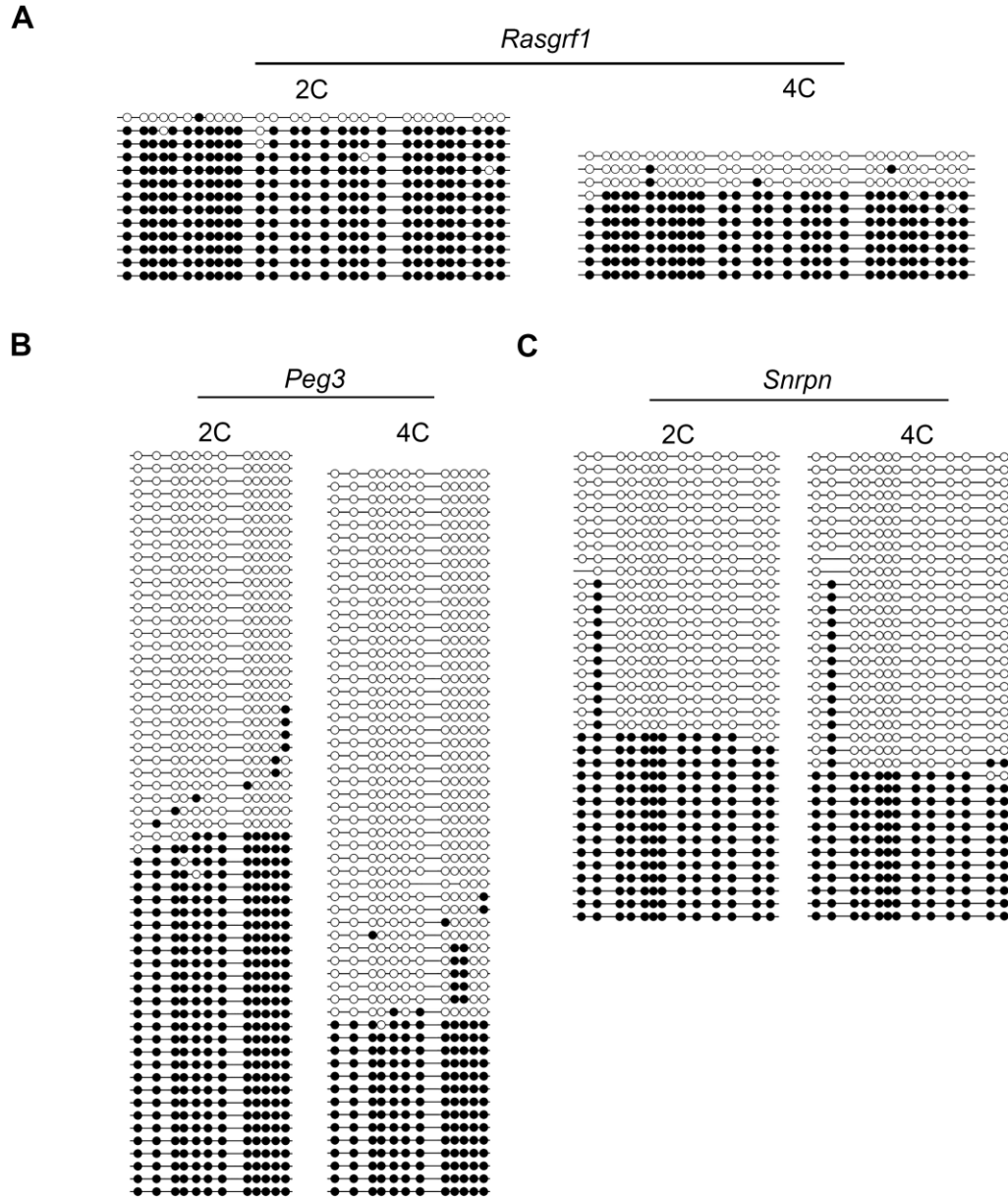


Figura 19. Análisis del estado de metilación de las gDMRs de *Rasgrf1*, *Peg3* y *Snrpn* mediante secuenciación de Sanger. Se muestra el estado de metilación de los clones secuenciados para *Rasgrf1* (A), *Peg3* (B) y *Snrpn* (C). Los sitios CpG no metilados se representan mediante círculos blancos mientras que los metilados se representan con círculos negros.

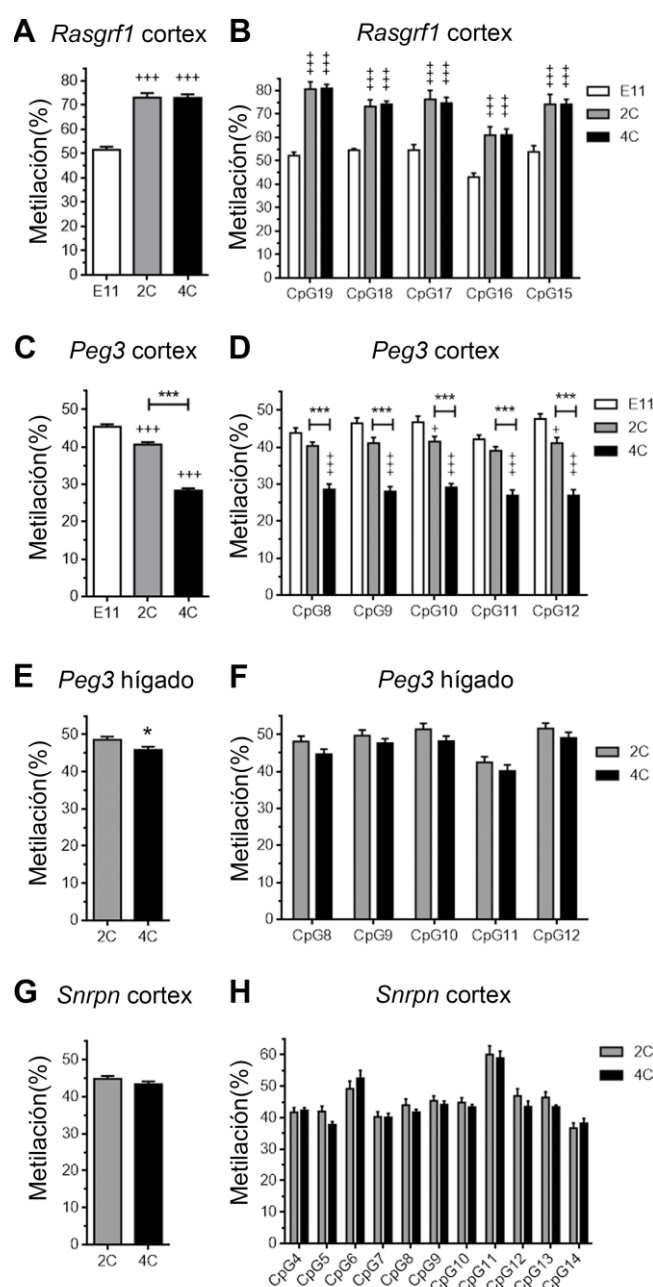


Figura 20. Aumento en los niveles de metilación de la Rasgrf1-DMR y disminución en los niveles de metilación de la Peg3-DMR de neuronas corticales (A) Análisis de pirosecuenciación global de los sitios CpG presentes en la Rasgrf1-DMR de células del neuroepitelio telencefálico de E11 y neuronas corticales diploides (2C) y tetraploides (4C). +++p-valor<0,005 (respecto E11) (B) Análisis de pirosecuenciación individual de sitios CpG (numerados como en **Figura 18**) presentes en la Rasgrf1-DMR de células del neuroepitelio telencefálico de E11 y neuronas corticales 2C y 4C. +++p-valor<0,005 (respecto E11) (C) Análisis de pirosecuenciación global de los sitios CpG presentes en la Peg3-DMR de células del neuroepitelio telencefálico de E11 y neuronas corticales 2C y 4C. +++p-valor<0,005 (respecto E11); ***p-valor<0,005 (respecto 2C). (D) Análisis de pirosecuenciación individual de sitios CpG (numerados como en **Figura 18**) presentes en la Peg3-DMR de células del neuroepitelio telencefálico de E11 y neuronas corticales 2C y 4C. +++p-valor<0,005; +p-valor < 0,05 (respecto E11); ***p-valor < 0,005 (respecto 2C). (E y G) Análisis de pirosecuenciación global de los sitios CpG presentes en la Peg3-DMR de hepatocitos 2C y 4C (E) y en la Snrpn-DMR de neuronas corticales 2C y 4C (G). *p-valor <0,05 (respecto 2C). (F y G) Análisis de pirosecuenciación individual de sitios CpG (numerados como en **Figura 18**) presentes en la Peg3-DMR de hepatocitos 2C y 4C (F) y en la Snrpn-DMR (numerados como **Figura 18**) de neuronas 2C y 4C (H).

4.7. Hemimetilación de Peg3 en subpoblación de neuronas corticales

Las técnicas mostradas en el apartado anterior, por su diseño de oligonucleótidos, sólo aportaban información sobre una de las dos hebras del ADN. Por lo tanto, para comprobar la existencia de hemimetilación en el dominio de Peg3 se empleó la técnica descrita por

Resultados

Laird y colaboradores (2004), que permite analizar de manera simultánea los niveles de metilación de las dos hebras que conforman una molécula de ADN.

El ADN obtenido a partir de núcleos neuronales diploides y tetraploides se trató con la enzima BbvI, de la cual se encontró una diana de corte dentro de la Peg3-DMR. A continuación, se le unió una horquilla de ADN previamente diseñada para este experimento (**Figura 21** y **Tabla 1**). Esta horquilla mantuvo unidas las dos hebras complementarias del ADN que seguidamente fueron linearizadas y convertidas con bisulfito. El ADN bisulfitado se amplificó por PCR con oligonucleótidos específicos *upstream* de la secuencia cortada (**Figura 21** y **Tabla 1**).

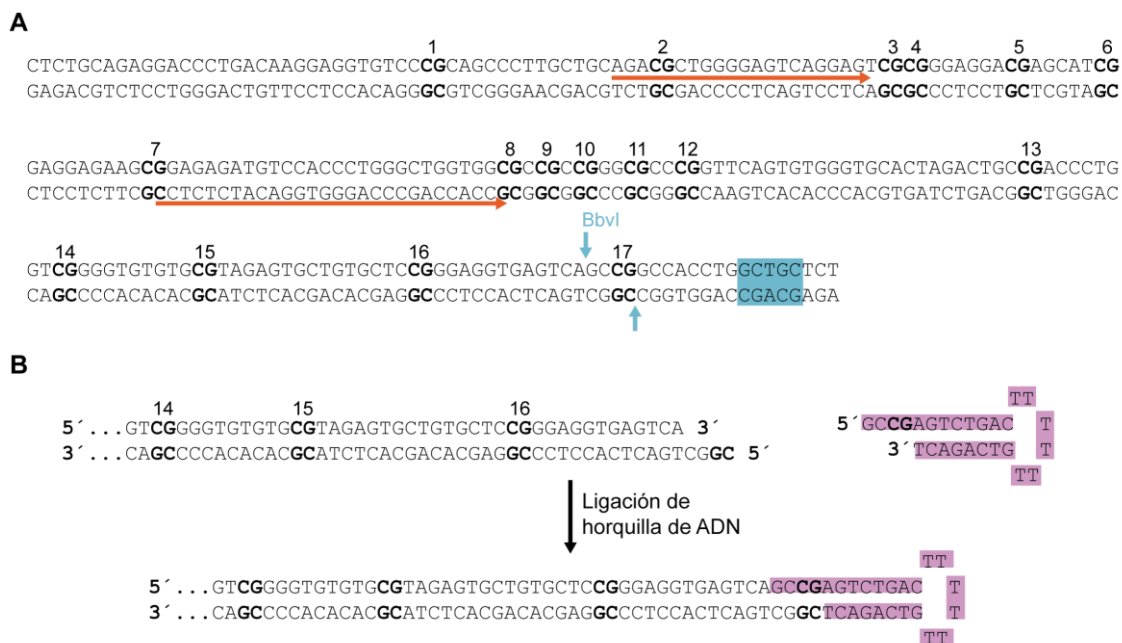


Figura 21. Esquema de la técnica *hairpin-bisulfite PCR* (A) Secuencia de ADN genómico de Peg3-DMR. Las CpGs aparecen numeradas y en negrita. La diana de restricción de BbvI se muestra en azul. Las flechas sólidas naranjas corresponden a los oligonucleótidos modificados (las Cs se sustituyen por Ts) utilizados en la PCR. (B) Diagrama que muestra el ADN cortado con BbvI y unido a la horquilla de ADN (lila). El ADN resultante de esta unión se trata con bisulfito y se amplifica con los oligonucleótidos indicados en (A). Las bandas resultantes de la amplificación se clonaron y se secuenciaron con secuenciación de Sanger (ver **Figura 22**).

Después de secuenciar con el método de Sanger los fragmentos de PCR, se analizó la presencia de Cs convertidas y no convertidas en ambas hebras de ADN (**Figura 22**). Este análisis demostró que, en neuronas tetraploides, el 15% de los clones secuenciados mostraba desmetilación hebra-específica en la Peg3-DMR (**Figura 22 A y C**). Además, se

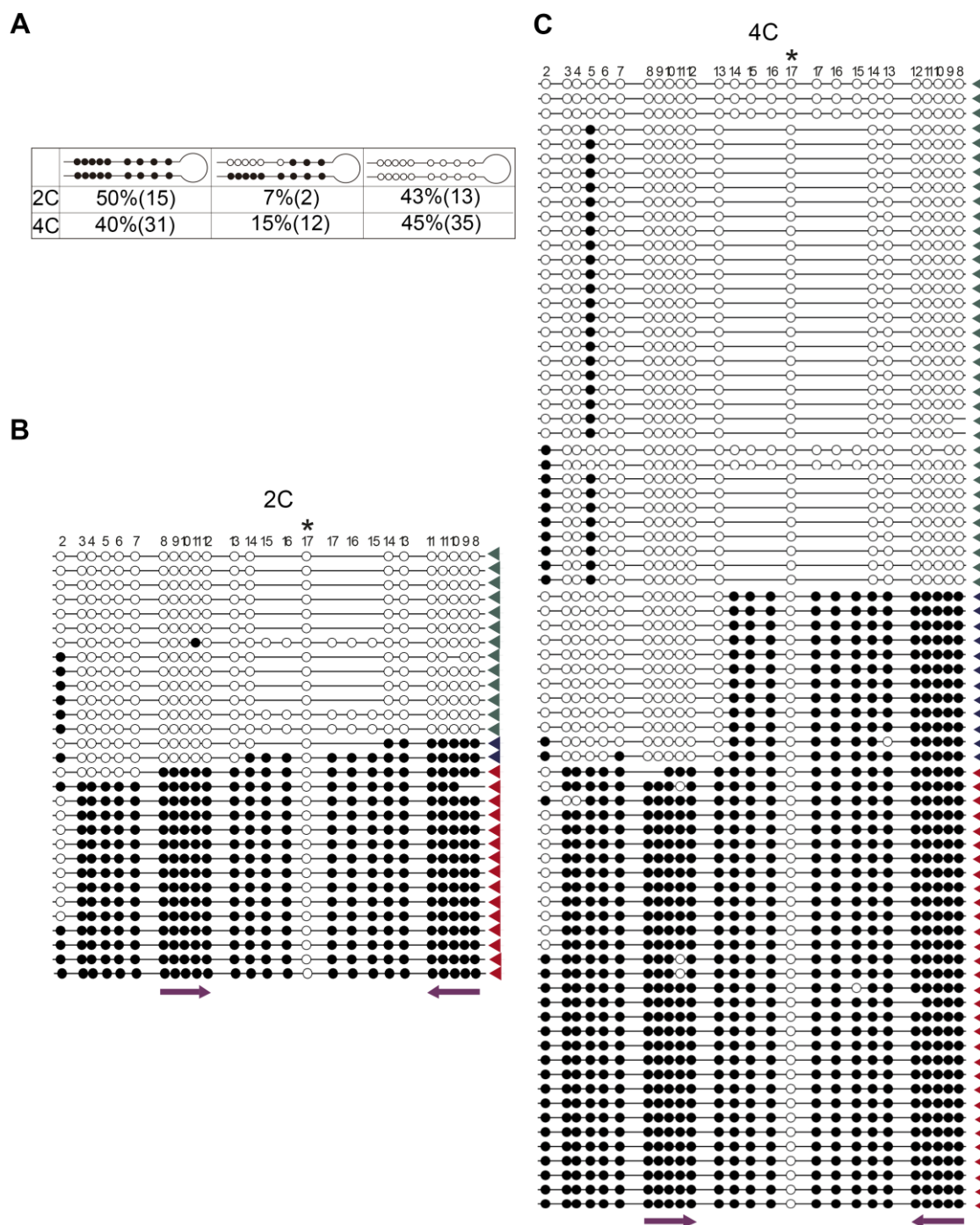


Figura 22. La reducción en los niveles de metilación de la Peg3-DMR de las neuronas corticales es hebra-específica (A) Porcentaje y número absoluto de clones obtenidos mediante la técnica de hairpin-bisulfite PCR [ver (B) y (C)] para los patrones de metilación indicados. Para todas las figuras, las CpG no metiladas se representan como círculos blancos mientras que las metiladas se representan como círculos negros. **(B y C)** Secuencias Sanger de los clones obtenidos por hairpin-bisulfite PCR a partir de ADN genómico derivado de neuronas diploides (2C) (B) y tetraploides (4C) (C). Triángulos verdes: clones con CpGs #8-12 (flechas moradas) no metiladas; triángulos azules: clones con CpGs #8-12 hemimetilados, triángulos rojos: clones con CpG #8-12 bimetilados. Los asteriscos indican la CpG #17, incluida en la horquilla (ver **Figura 21 B**), que por lo tanto estará no metilada en todos los casos. La CpG #2, que forma parte del oligonucleótido empleado para amplificar, estaba degenerada.

Resultados

observó una pequeña proporción de clones con desmetilación hebra-específica en la Peg3-DMR de neuronas diploides (7%) (**Figura 22 A y B**), en consonancia con los resultados del WGBS y pirosecuenciación anteriores (**Figura 16, Figura 17 y Figura 20**). En ambos tipos neuronales, la desmetilación afectaba a las CpGs localizadas en la hebra negativa de Peg3 (**Figura 16 y 17**).

Con todo esto se concluyó que una proporción de CpGs de la Peg3-DMR del cromosoma metilado, comprendiendo al menos las CpGs #8 a 13 (**Figura 22 B y C**) que coincidían además con las CpG hemimetiladas detectadas con el WGBS, se encuentra asimétricamente metilada en neuronas corticales de P15.

No pudimos encontrar ningún SNP en esta región que nos permitiese determinar si la hemimetilación se daba en el cromosoma paterno o materno. Pero este análisis mostró que las CpGs #14 a 17 se encontraban metiladas en ambas hebras (**Figura 22 B y C**) en la mitad de los clones y no metiladas en la otra mitad, como se espera de la metilación diferencial entre los cromosomas materno y paterno, respectivamente, asociada al imprinting genómico. Dado que los clones hemimetilados estaban todos bimetilados en las CpGs #14 a 17, suponemos que la disminución en la metilación de las neuronas corticales observada (**Figura 20 C y D**) se debe a la eliminación asimétrica de 5mCs en la Peg3-DMR del cromosoma materno.

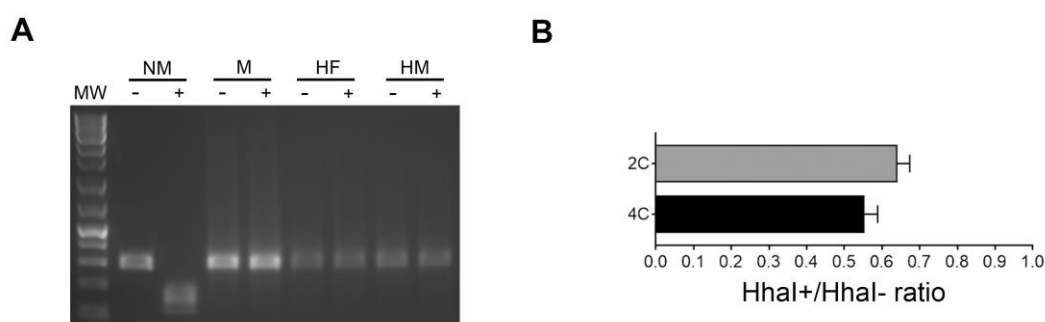


Figura 23. Confirmación de hemimetilación en la Peg3-DMR mediante corte con enzima HhaI (A) Ensayo de restricción con la enzima HhaI empleando ADN no metilado (NM), bimetilado (M), formilado en una hebra y metilado en la complementaria (HF), y ADN hemimetilado (HM). (-): ADN sin HhaI, (+): ADN tratado con HhaI, MW: marcador de peso molecular para ADN. **(B)** ADN genómico de núcleos neuronales de diploides (2C) y tetraploides (4C), no digerido (HhaI-) o digerido con HhaI (HhaI+), amplificado por qPCR empleando oligonucleótidos especificados en Figura 18). Se representa la media de la ratio HhaI+/HhaI- (\pm s.e.m.). Los valores no eran significativamente menores a 0,5 (analizado con test de t-Student de una cola).

La hemimetilación descrita anteriormente se confirmó también mediante un ensayo con una enzima de restricción seguido de PCR cuantitativa (Hashimoto *et al.*, 2007; Hua *et al.*, 2011; Bai *et al.*, 2015). Para este análisis se eligió la enzima HhaI, ya que esta no corta el ADN bimetilado, hemimetilado (Bird, 1978), ni el metilformilado (**Figura 23 A**). Se aisló ADN genómico de núcleos neuronales diploides y tetraploides que se digirió con la enzima HhaI. El ADN se amplificó a continuación mediante PCR cuantitativa utilizando oligonucleótidos que flanqueaban los sitios de corte de la enzima HhaI dentro de la Peg3-DMR [CpGs #8 y #11 (**Figura 18**)].

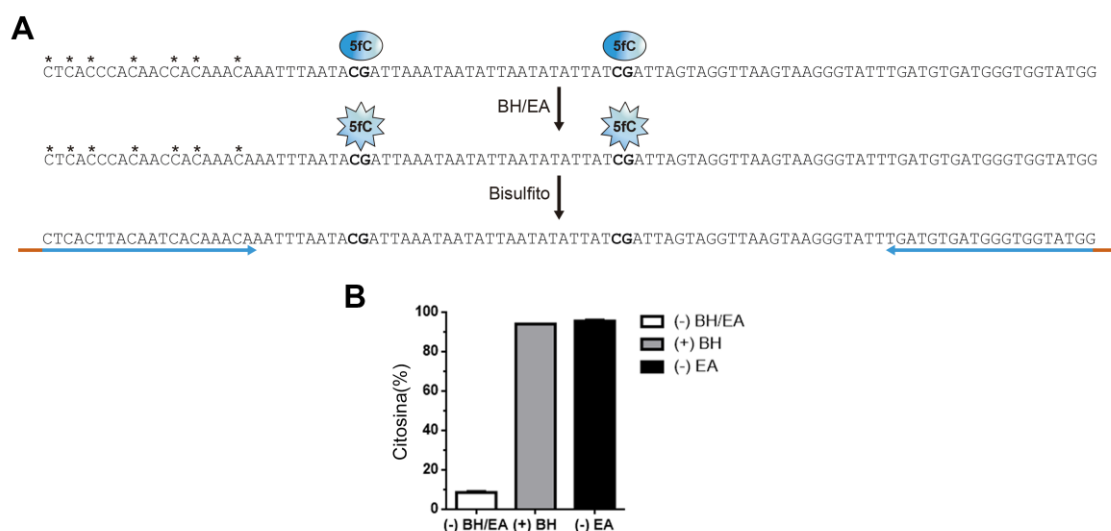
Este ensayo demostró que aproximadamente el 50% del ADN genómico permanecía sin cortar tanto en neuronas diploides como en tetraploides (**Figura 23 B**), indicando que la mitad de los cromosomas eran resistentes al corte con HhaI en la Peg3-DMR. Estos resultados, junto con los resultados de la pirosecuenciación, en los que se veía una disminución en el porcentaje de metilación (**Figura 20 C y D**), sólo son compatibles con una desmetilación parcial hebra-específica de ciertos sitios dentro de la Peg3-DMR.

Por tanto, mediante tres técnicas independientes, WGBS, hairpin-bisulfite PCR y corte con enzima de restricción sensible a metilación seguida de qPCR, se demostró la existencia de metilación hebra-específica en la Peg3-DMR.

4.8. La reducción en la metilación de la Peg3-DMR de neuronas corticales es independiente de la presencia de 5fC

La 5fC es una modificación estable del ADN que se encuentra en el genoma de mamíferos (Bachman *et al.*, 2015) pero que no se detecta mediante la técnica clásica de tratamiento con bisulfito seguido de secuenciación (Plongthongkum *et al.*, 2014), leyéndose como una T (**Figura 4**). Para descartar que la hemimetilación observada en los sitios CpG dentro de la Peg3-DMR contuviese 5fC, se llevó a cabo un ensayo con BH y con EA. Estos compuestos modifican específicamente la 5fC e impiden que sea transformada por el tratamiento con bisulfito (Song *et al.*, 2013; Plongthongkum *et al.*, 2014), leyéndose como una C tras la secuenciación (**Figura 4**).

En primer lugar, evaluamos la capacidad protectora frente al tratamiento con bisulfito del BH y de la EA. Para ello sintetizamos una sonda de ADN que contenía 5fC en dos CpGs específicas (**Figura 24 A**). Dicha sonda se trató con BH o con EA y se convirtió con bisulfito. A continuación, se generaron las librerías para la NGS, con los adaptadores



diploides (**Figura 25 G y H**) y hepatocitos tetraploides (**Figura 25 I y J**). Por último, no se vieron cambios significativos cuando se realizaron los mismos análisis en ADN genómico de neuronas diploides utilizando primers específicos para la *Snrpn*-DMR, ni globales (**Figura 25 K**) ni en los sitios CpG analizados individualmente (**Figura 25 L**).

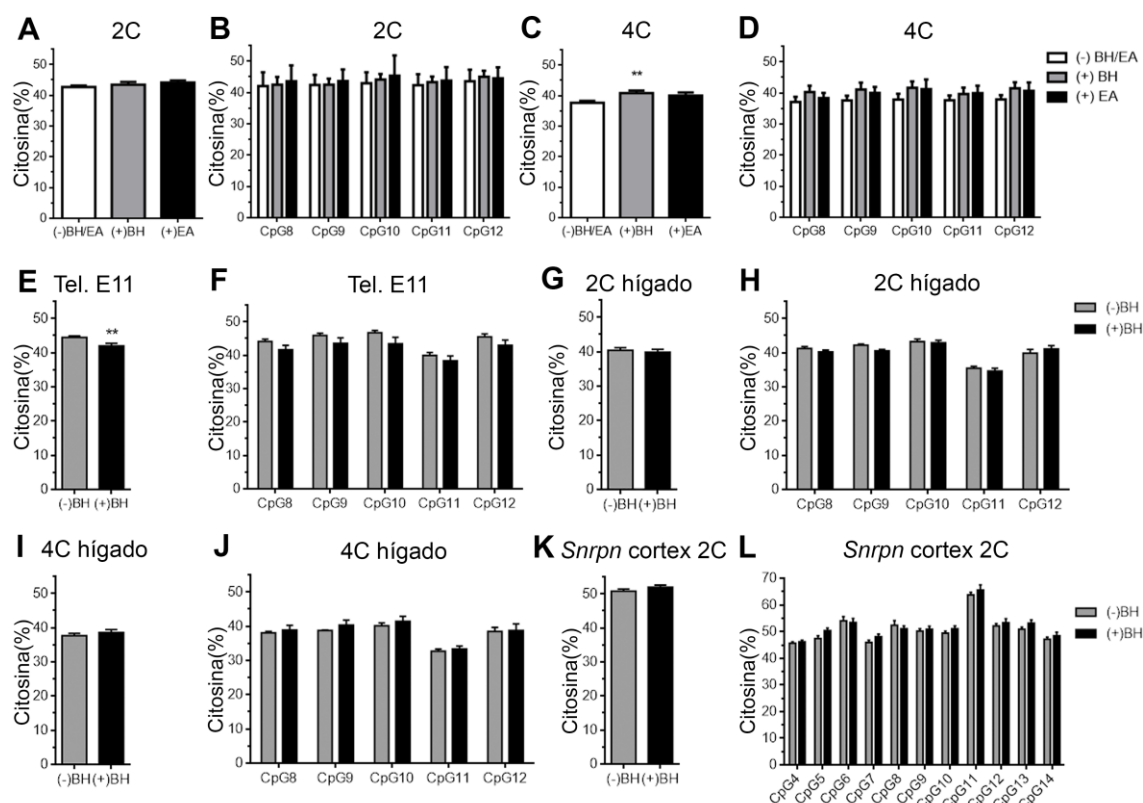


Figura 25. La reducción de los niveles de metilación en la Peg3-DMR de neuronas corticales tetraploides (4C) podría ser parcialmente debida a la presencia de 5fC. (A) Análisis de pirosecuenciación global de los sitios CpG presentes en la Peg3-DMR, realizados con ADN genómico no tratado (blanco), tratado con BH (gris) o tratado con EA (negro) de neuronas corticales diploides (2C). (B) Análisis de pirosecuenciación individual de los sitios CpG (numerados como en **Figura 18 B**) realizados con ADN genómico no tratado (blanco), tratado con BH (gris) o tratado con EA (negro) de neuronas corticales 2C. (C) Análisis de pirosecuenciación global de los sitios CpG presentes en la Peg3-DMR, realizados con ADN genómico no tratado (blanco), tratado con BH (gris) o tratado con EA (negro) de neuronas corticales 4C. (D) Análisis de pirosecuenciación individual de los sitios CpG (numerados como en **Figura 18 B**), realizados con ADN genómico no tratado (blanco), tratado con BH (gris) o tratado con EA (negro) de neuronas corticales 4C. (E, G e I) Análisis de pirosecuenciación global de los sitios CpG presentes en la Peg3-DMR, realizados con ADN genómico no tratado (gris) o tratado con BH (negro) de neuroepitelio telencefálico de E11 (E), hepatocitos 2C (G) y hepatocitos 4C (I). (F, H y J) Análisis de pirosecuenciación individual de los sitios CpG (numerados como en **Figura 18 B**), realizados con ADN genómico no tratado (gris) o tratado con BH (negro) de neuroepitelio telencefálico de E11 (F), hepatocitos 2C (H) y hepatocitos 4C (J). (K) Análisis de pirosecuenciación global de los sitios CpG presentes en la *Snrpn*-DMR, realizados con ADN genómico no tratado (gris) o tratado con BH (negro) de neuronas corticales 2C. (L) Análisis de pirosecuenciación individual de los sitios CpG (numerados como en la **Figura 18 C**), realizados con ADN genómico no tratado (gris) o tratado con BH (negro) de neuronas corticales 2C. **p-valor < 0.01.

Resultados

Estos resultados sugieren que la *Peg3*-DMR contiene CpGs diferencialmente metilformiladas en una subpoblación de neuronas corticales tetraploides. Si tenemos en cuenta los resultados anteriores (**Figura 22 y 23**), este patrón, que representa una marca epigenética nunca antes descrita, parece derivar de una modificación química de una C de una de las dos hebras del ADN.

*4.9. Mantenimiento del imprinting del dominio *Peg3* en las neuronas corticales*

Para estudiar si la hemimetilación de las CpGs participa en el mantenimiento del imprinting genómico del dominio *Peg3*, nos centramos en el gen *Peg3*. Se analizó el silenciamiento del alelo materno de dicho gen empleando un SNP (# rs31442449) localizado en el intrón 7 del gen *Peg3* (**Figura 26 A**). Este SNP está caracterizado por un cambio de T en ratones C57BL/6J a C en ratones DBA/2J. Ratones hembra C57BL/6J se cruzaron con ratones machos DBA/2J para generar ratones heterocigotos híbridos para el polimorfismo T/C en los cromosomas materno/paterno, respectivamente (**Figura 26 B**).

Se obtuvo ARNm naciente (enriquecido en transcritos que no han sufrido *splicing*) a partir de núcleos neuronales diploides y tetraploides, separados mediante FACS a partir de corteza cerebral de ratones heterocigotos P15. Además, se obtuvo ARNm de núcleos de células del neuroepitelio telencefálico de embriones E11 heterocigotos. Se comprobó que estos ARNm no contenían ADN genómico contaminante tras amplificar y reamplificar (100 ciclos de PCR en total) con los oligonucleótidos empleados en el análisis posterior sin obtener ningún amplificación (**Figura 26 C**). Además, los resultados indicando la ausencia de expresión del alelo materno (**Figura 26 y 27**) descartan la contaminación con ADN genómico.

El análisis con pirosecuenciación de los ADNc obtenidos a partir de los ARNm anteriormente descritos, indicaron que el alelo materno del gen de *Peg3* se mantiene silenciado, ya que sólo se obtuvo señal de secuencias con T tanto en neuronas corticales diploides (**Figura 26 E**) como en tetraploides (**Figura 26 F**), similar a lo observado en el neuroepitelio de E11 (**Figura 26 D**). Adicionalmente, el silenciamiento del alelo materno del gen *Peg3* en el neuroepitelio de E11 y en neuronas 2C se confirmó utilizando oligonucleótidos adicionales más cercanos al SNP (**Figura 26 G y H**).

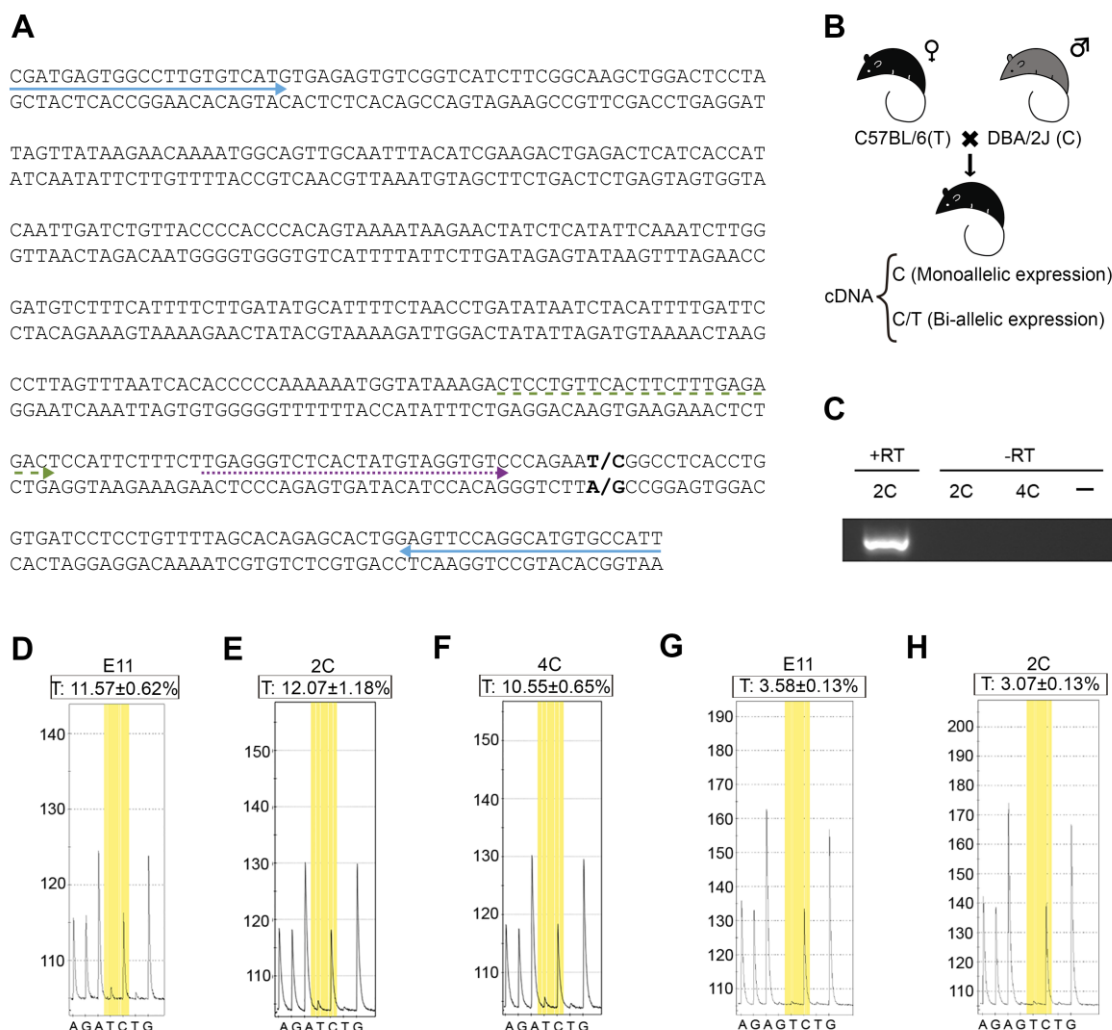


Figura 26. La hemimetilación de Peg3-DMR mantiene la expresión monoalélica del gen *Peg3*.

(A) Muestra la secuencia parcial del ADNc a partir del transcrito no *spliceado* (amplificado utilizando los oligonucleótidos indicados con flechas sólidas). Primers específicos [flecha discontinua verde: oligonucleótidos utilizados para (D-F); flecha punteada morada: oligonucleótidos empleados para (G y H)] se emplearon para pirosecuenciar la región que contenía el SNP # rs31442449 (señalado en negrita). (B) Esquema de los cruces de ratones con el SNP en el gen *Peg3*, y los ADNc esperados en el caso de expresión mono- o bialélica. (C) Amplificación por PCR con los oligonucleótidos indicados con flechas sólidas en (A) usando un control negativo para RT (-RT) o ARNm de núcleos neuronales de diploides (2C) o tetraploides (4C) tratados con RT (+RT) o agua (-). Al ADNc +RT se le aplicaron 50 ciclos de PCR y a la reacción -RT se le aplicaron 50+50 ciclos de PCR. (D-F) Pirogramas representativos de la secuencia de ADNc de neuroepitelio telencefálico de E11 (D), neuronas corticales 2C (E) y 4C (F), empleando los oligonucleótidos indicados con la flecha discontinua verde en (A). El SNP (T/C) se muestra en amarillo. La T que aparece en la posición 6 es un control negativo. En cada análisis, se emplearon dos ADNc. Se muestra la media de la proporción de T en la posición 4 (T4) (\pm s.e.m.). (G y H) Pirogramas representativos de la secuencia de ADNc de neuroepitelio telencefálico de E11 (G) y neuronas corticales 2C (H), empleando los oligonucleótidos indicados con la flecha punteada morada en (A). El SNP (T/C) se muestra en gris. La G de la posición 4 (G4) y la T de la posición 7 (T7) son controles negativos. Se muestra la media de la proporción de T en la posición 5 (T5) (\pm s.e.m.).

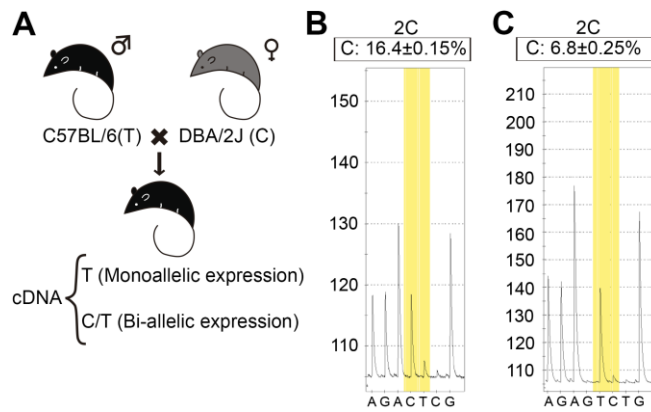


Figura 27. La hemimetilación de la Peg3-DMR es suficiente para mantener la expresión monoalélica del gen *Peg3*. (A) Esquema del cruce recíproco al descrito en la Figura 25. (B,C) Pirogramas representativos del cDNA pirosecuenciado de neuronas corticales diploides (2C) empleando el oligonucleótido señalado como una flecha discontinua (B) o una flecha punteada (C) en la Figura 25. El SNP (T/C) se muestra en amarillo. La T6 en (B) y la G4 y la T7 en (C) son controles negativos. Se muestra la media de la proporción de C5 (B) y C6 (C) (\pm s.e.m.).

5. Discusión

Las neuronas tetraploides somáticas se generan mediante un proceso perfectamente regulado, como se ha demostrado en estudios anteriores del laboratorio (López-Sánchez y Frade, 2013; López-Sánchez *et al.*, 2011; Morillo *et al.*, 2010), y se mantienen a lo largo de la vida adulta del individuo con el genoma completamente duplicado, tal y como se ha visto en esta tesis (**Figura 7 B**).

En organismos con poliploidía germinal, como plantas (Madlung y Wendel, 2013), se ha descubierto que la poliploidización da lugar a cambios epigenéticos (Chen, 2007; Song y Chen, 2015). La poliploidización y el mantenimiento de mecanismos de regulación epigenética, como el imprinting genómico, también se había abordado en mamíferos con tetraploidía germinal (Bacquet *et al.*, 2008; Yamazaki *et al.*, 2016). En nuestro caso, quisimos verificar la existencia de mecanismos implicados en el establecimiento de marcas epigenéticas, como la metilación, en células con tetraploidía somática. Para ello, se llevó a cabo un análisis global, caracterizando el estado de metilación del genoma completo de las neuronas tetraploides. Se eligió el estadio P15, en el que sólo hay neuronas tetraploides recién diferenciadas, para evitar la presencia de neuronas tetraploides generadas *de novo* en el cerebro adulto (López-Sánchez *et al.*, 2017)

Uno de los principales problemas a la hora de realizar este estudio fue el reducido porcentaje de neuronas tetraploides presente en el cortex de ratón, que complicó tanto la purificación de los núcleos tetraploides mediante FACS, como las técnicas de análisis posteriores, entre ellas, el análisis por WGBS. Hasta hace poco era necesaria una gran cantidad de ADN para llevar a cabo esta técnica, ya que el tratamiento con bisulfito, degradaba gran parte de éste. El kit comercial empleado en estos ensayos permitía trabajar con cantidades muy reducidas de ADN de partida, incluyendo un paso previo de amplificación con random primers específicos para el ADN bisulfitado, lo que nos permitió generar las librerías para realizar WGBS con las neuronas tetraploides.

Este kit, sin embargo, no incluía ninguna sonda control que permitiese evaluar la conversión de bisulfito. Se recurrió a otra forma para determinar la eficiencia de la conversión, que consistía en chequear los sitios no-CpG, ya que un alto número de ellos es indicativo de una conversión incompleta. El tejido nervioso adulto es precisamente un caso particular, donde se ha detectado metilación en sitios no-CpG (Xie *et al.*, 2012), pero

comparando el número de no-CpGs metiladas detectadas en nuestra muestra (TETRA) y en la muestra FCX (extraída de Lister *et al.*, 2013), determinamos que la conversión había sido adecuada, dado que el número en ambas muestras era similar (**Tabla 2**).

A pesar de que el número de lecturas obtenido en la secuenciación de las dos muestras procedentes de neuronas tetraploides no era tan diferente del obtenido en la secuenciación llevada a cabo por Lister y colaboradores (2013) (**Tabla 2**), el alineador empleado para el mapeo de dichas lecturas con el genoma de referencia descartó muchas de ellas, reduciendo el porcentaje de lecturas mapeadas (**Tabla 2**). Afortunadamente, al tener dos réplicas reproducibles de la misma muestra y combinarlas, finalmente pudimos secuenciar aproximadamente los mismos sitios CpGs y con una cobertura promedio equiparable en las muestras TETRA y FCX para los siguientes análisis (**Figura 8**).

Al analizar los patrones de metilación global en neuronas tetraploides y células de FCX, se vio que en ambas muestras predominaba la hipermetilación en sitios CpG (**Figura 9 A**), lo que cuadra con observaciones previas de otros tejidos adultos y del propio tejido nervioso de ratón (Hon *et al.*, 2013; Lister *et al.*, 2013). El porcentaje de CpGs parcialmente metiladas también estaba en consonancia con lo encontrado en estudios anteriores en los que se emplearon regiones completas del cerebro en vez de neuronas separadas mediante FACS (Hon *et al.*, 2013). Algunos autores (Kulis *et al.*, 2015) achacan este alto porcentaje de CpGs parcialmente metiladas a trabajar con poblaciones celulares mixtas, sin separar mediante FACS. Es posible que este sea el motivo por el que este porcentaje de CpGs parcialmente metiladas es mayor en la muestra de FCX, en la que se incluyen tanto neuronas como células de glía, que en la muestra TETRA, en la que sólo hay neuronas tetraploides. Probablemente, el hecho de que el porcentaje de CpGs parcialmente metiladas sea elevado en esta última muestra, se deba también a que, dentro de las neuronas tetraploides aisladas podemos encontrar diferentes tipos neuronales, ya que la separación sólo se hizo en base a su marcaje con NeuN, y sin ningún otro marcador más específico de tipo neuronal.

Los datos de metilación de los sitios no-CpG también concordaban con los análisis previos (Lister *et al.*, 2013; Guo *et al.*, 2014), con niveles de metilación en torno al 30%, CAC como contexto más frecuente y los menores niveles de metilación en el cromosoma Y (**Figura 11**). El número de posiciones no-CpG metiladas detectadas con nuestro análisis que coincidía entre las muestras FCX y TETRA era muy pequeño (datos no mostrados). El

análisis de la distribución de los niveles de metilación de estos sitios, sin embargo, nos permitió demostrar que en las neuronas tetraploides hay más metilación en sitios no-CpG que en las células del cortex frontal. Esto puede deberse una vez más a que la muestra de cortex frontal incluye también glía, la cuál presenta menores niveles de metilación en estos sitios (Lister *et al.*, 2013) y podría estar reduciendo los niveles de metilación de no-CpGs en la muestra PCX. Lister y colaboradores (2013) también observaron que los niveles de metilación de no-CpGs aumentaba durante el desarrollo y coincidiendo con la sinaptogénesis y el aumento de la densidad sináptica. Teniendo en cuenta esto, el aumento en los niveles de metilación de no-CpGs observado en las neuronas tetraploides podría deberse a que el proceso de desarrollo esté más acelerado en estas neuronas. El hecho de que los sitios no-CpG detectados simultáneamente en la muestra PCX y TETRA fuese tan pequeño (datos no mostrados) era algo esperable, ya que la metilación en estos sitios es un fenómeno altamente variable (Ziller *et al.*, 2011).

La identificación de dominios con diferentes niveles de metilación entre las muestras analizadas mediante WGBS es importante ya que se piensa que estos DMDs pueden regular la expresión génica en función del tipo celular (Deaton y Bird, 2011). Existen diferentes paquetes informáticos con distintas aproximaciones estadísticas para detectar DMDs (revisados en Bock, 2012), pero estos suelen requerir réplicas biológicas para estimar la varianza dentro de los grupos. En este estudio, sólo disponíamos de una réplica por grupo, por lo que tuvimos que recurrir a la librería DSS (Wu *et al.*, 2015) para R. Se trata de un método basado en el test de Wald que puede ser utilizado cuando sólo existe una réplica por grupo. Dado que los niveles de metilación de los sitios CpGs cercanos son similares, estos pueden considerarse como “pseudo-réplicas” y estimarse con ellos la dispersión necesaria para el test de Wald.

Las DMDs identificadas con este método se anotaron funcionalmente empleando herramientas bioinformáticas, lo que arrojó algunos resultados interesantes. En general se vio que la mayoría de las DMDs, se asociaban con procesos biológicos o compartimentos celulares de neuronas. Específicamente, se vio que muchos genes que presentaban DMDs estaban implicados en procesos biológicos como el desarrollo y morfogénesis de proyecciones neuronales, morfogénesis o sinapsis. En estudios anteriores (López-Sánchez y Frade, 2013) se ha demostrado que un elevado porcentaje de las neuronas tetraploides de corteza (60-85%), expresan el factor de transcripción específico de neuronas de proyección

CTIP2 (Arlotta *et al.*, 2005). También se ha visto que las neuronas tetraploides presentan un tamaño superior a las neuronas diploides (Morillo *et al.*, 2010; Sigl-Glöckner y Brecht, 2017). Es probable que las diferencias en la metilación entre la muestra FCX y la muestra TETRA se correspondan con un patrón epigenético específico que se traduzca en las particularidades fenotípicas de las neuronas tetraploides, como mayor tamaño o largas proyecciones.

Se trata de resultados preliminares, que deben ser confirmados mediante otras técnicas experimentales. En primer lugar, habría que confirmar la reproducibilidad biológica de las DMDs mediante ensayos específicos como pirosecuenciación de las regiones concretas. En segundo lugar, y más importante, habría que acompañar el análisis mediante WGBS de un análisis ARN-Seq que permita ver si realmente esas variaciones en la metilación de ciertas regiones afectan a la expresión de los genes en las que se encuentran, como se ha hecho en otros estudios (Lister *et al.*, 2009, Lister *et al.*, 2013). Este tipo de análisis es especialmente complicado en el caso de las neuronas tetraploides. Al utilizar núcleos neuronales para el FACS, perdemos el ARN citoplasmático y de los terminales sinápticos, no siendo capaces de obtener una cantidad de ARN suficiente para llevar a cabo un ARN-Seq. En el futuro, empleando técnicas de *single cell* de reciente desarrollo (Sheng *et al.*, 2017), esperamos ser capaces de llevar a cabo este análisis.

Respecto al imprinting genómico en las neuronas tetraploides, con los datos obtenidos por WGBS pudimos ver que los patrones de metilación globales de las gDMRs que controlan genes de imprinting se asemejaban a los descritos en la literatura para estas regiones (**Figuras 14 y 15**). Tan sólo la gDMR de *Dlk1-Gtl2* presentaba niveles de metilación cercanos al 100%. Sin embargo, si en vez de emplear las coordenadas de WAMIDEX que cubren una mayor extensión (ver Materiales y métodos), restringimos la gDMR de *Dlk1-Gtl2* a las coordenadas descritas por otros autores (posición 109528345-109528471, Lin *et al.*, 2013; Guntrum *et al.*, 2017), se puede comprobar que el valor de metilación de dicha gDMR sí se aproxima al 50% descrito en la literatura para tejidos con expresión monoalélica de *Dlk1* (Takada *et al.*, 2002; Ferrón *et al.*, 2011).

Los datos obtenidos mediante WGBS, que contenían información del porcentaje de metilación de la hebra positiva y negativa del ADN, nos permitieron detectar metilaciones hebra-específicas en las CpGs de muchas gDMRs incluidas las de *Peg3* y *Rasgrfl* (**Figuras 16 y 17**). Consideramos este resultado altamente relevante pues cuestiona la

invariabilidad del patrón de metilación de las gDMRs una vez adquirido durante el desarrollo embrionario, como ya apuntaban otros estudios previos en los que se detectaron cambios en los patrones de metilación en contextos no patológicos (Ferrón *et al.*, 2011).

Pudimos confirmar la hemimetilación hebra-específica en los dominios de imprinting de los genes *Peg3* y *Rasgrfl* mediante otras dos técnicas: bisulfito seguido de secuenciación Sanger y bisulfito seguido de pirosecuenciación. Cuando se llevan a cabo estas técnicas, es necesario tener en cuenta que la conversión con bisulfito va a provocar que las dos hebras de ADN pierdan su complementariedad (**Figura 28**). Esto es algo a considerar a la hora de diseñar los oligonucleótidos para la PCR previa a la clonación y a la pirosecuenciación. En nuestro caso, el diseño de oligonucleótidos se hizo basándonos en los datos de WGBS, para analizar el estado de metilación de la hebra hipermetilada en el caso de *Rasgrfl* y de la hebra que presentaba una disminución en la metilación en el caso de *Peg3*. Para *Snrpn*, se eligió una de las dos hebras, sin tener en cuenta cual, pues el análisis mediante WGBS indicaba que la metilación iba a ser igual en ambas.



Figura 28. La elección de oligonucleótidos es esencial a la hora de realizar el análisis de la metilación con métodos basados en la conversión con bisulfito. Las dos hebras del ADN pierden su complementariedad después de la conversión con bisulfito. El oligonucleótido indicado con una flecha morada amplificará la hebra 5'-3' generando una nueva hebra cuya secuencia no coincide con la de la hebra 3'-5' inicial. Esta nueva hebra será a su vez amplificada con el oligonucleótido complementario al indicado con una flecha verde. De esta forma, sólo obtendremos información sobre el estado de metilación de la C de la hebra 5'-3' inicial.

En el caso concreto de *Peg3* se emplearon además técnicas como el corte de ADN con una enzima sensible a metilación seguida de PCR a tiempo real y hairpin-bisulfite PCR para confirmar la existencia de esta hemimetilación.

La hemimetilación hebra-específica no es un fenómeno específico de neuronas tetraploides, ya que también se detectó en las neuronas diploides y en el análisis de los datos del WGBS de FCX. Además, no todas las neuronas tetraploides mostraron esta modificación, si bien es verdad que este fenómeno apareció en mayor proporción en este tipo neuronal. Lo más probable es que esta marca epigenética afecte a un subtipo particular de neuronas corticales que sean principalmente tetraploides.

El dogma actual afirma que la metilación de la C en el contexto CG es siempre simétrico, y se mantiene así gracias a la acción de las DNMTs, que garantizan su propagación simétrica durante las divisiones celulares, asegurando la heredabilidad de esta marca epigenética (Bird, 2011; Chen y Riggs, 2011; Goll y Bestor, 2005; Jaenisch y Bird, 2003; Wigler *et al.*, 1981). El presente estudio (publicado en Patiño-Parrado *et al.*, 2017) pone en tela de juicio este dogma ya que, como hemos visto, estas gDMRs contienen subdominios hemimetilados y probablemente metilformilados.

Además, en el caso de la gDMR del dominio *Dlk1-Gtl2*, nuestros datos coinciden con los del reciente estudio publicado por Guntrum y colaboradores (2017). En este estudio se analizaba la existencia de hemimetilación en la IG-DMR, situada entre los genes *Dlk1* y *Gtl2*. Los autores, sin embargo, afirman que la hemimetilación, aunque presente en gDMRs (las analizadas en este estudio), es un fenómeno mucho más frecuente en sDMR.

Las hemimetilaciones en las gDMRs detectadas en esta tesis afectan a una de las dos hebras del ADN, y siempre a la misma. Esto sugiere la existencia de algún mecanismo molecular específico que metile, en el caso de *Rasgrf1*, o desmetile, en el caso de *Peg3*, las CpGs de una de las hebras del cromosoma materno.

Como hemos visto, la *Rasgrf*-DMR de neuronas de ratones P15, se encuentra hemimetilada y lo más probable es que sea la hebra positiva del cromosoma materno la que se encuentra metilada. *Rasgrf1* presenta expresión monoalélica del alelo paterno metilado en el cerebro de ratón únicamente en periodos neonatales. Después del destete, la expresión en el cerebro pasa a ser bialélica (Drake *et al.*, 2009). En estudios anteriores se ha especulado con la posibilidad de que este cambio se deba a la aparición de nuevos *enhancers* que se salten el bloqueo de la transcripción que ejerce la unión del CTCF al alelo materno (Yoon *et al.*, 2005). Hark y colaboradores (2000), demostraron que la unión de CTCF a la DMR de *H19-Igf2*, quedada inhibida cuando el ADN estaba metilado en la hebra positiva. Podría

ocurrir que la unión de CTCF al alelo materno de la Rasgrf-DMR se bloquee por la hemimetilación que hemos encontrado. Sería necesario realizar experimentos que confirmasen esta hipótesis.

La gDMR que controla el imprinting del dominio Peg3 se encuentra situada en el promotor del gen *Peg3*. Un estudio reciente de Perera y Kim (2016) ha demostrado que la delección del promotor materno de *Peg3* da lugar a una sobre-expresión del promotor paterno del mismo gen, lo que revela una compleja regulación en este dominio. Como ocurre en el caso de la formilación (Ngo *et al.*, 2016) y de la hemiformilación (Raiber *et al.*, 2015) de CpG, es posible que la metilformilación observada en este estudio altere la estructura del promotor materno de *Peg3*, que podría dar lugar a cambios en los niveles de expresión de *Peg3* a partir del promotor paterno, sin modificar la expresión monoalélica. No puede descartarse tampoco la posibilidad de que esta 5fC interaccione con proteínas con funciones específicas (Song y Pfeifer, 2016).

Como se indicó en la *Introducción*, todavía no se ha descrito ninguna desmetilasa capaz de llevar a cabo una desmetilación directa, rompiendo la unión covalente entre el metilo y la C. La presencia de 5fC detectada en neuronas tetraploides, en las CpG # 8 a 12 de la gDMR de *Peg3*, podría interpretarse como un estado intermedio en la desmetilación hebra-específica antes de la BER (Maiti y Drohat, 2011). La timina ADN glucosilasa (TDG) puede reconocer y escindir selectivamente 5fC y 5caC, desencadenando el proceso de reparación que da lugar a una C sin modificar (He *et al.*, 2011; Neri *et al.*, 2015). Se requieren más investigaciones para tratar de elucidar el mecanismo molecular implicado en la generación de esta hemimetilación hebra-específica y para comprender la función de esta nueva marca epigenética.

6. Conclusiones

1. El genoma de las neuronas tetraploides corticales se encuentra completamente duplicado.
2. El patrón de metilación en CpGs de las neuronas tetraploides corticales es similar al del resto de células corticales, aunque existen DMDs en genes específicos que podrían dar lugar a las particularidades fenotípicas de las neuronas tetraploides.
3. Los niveles de metilación en sitios no-CpGs en neuronas tetraploides corticales son mayores que en el resto de células corticales. Esto podría deberse a una aceleración en el proceso de maduración de las neuronas tetraploides.
4. Existe hemimetilación en gDMRs de genes sujetos a imprinting, tanto en neuronas tetraploides como diploides en la corteza.
5. La Peg3-DMR contiene CpGs diferencialmente metil-formiladas en una subpoblación de neuronas tetraploides corticales.
6. La hemimetilación de Peg3-DMR mantiene la expresión monoalélica del gen *Peg3*.
7. Dado que las hemimetilaciones en las gDMRs detectadas son hebra-específicas, es bastante probable que exista un mecanismo molecular específico que de lugar a este tipo de modificaciones.

7. Bibliografía

- Arlotta, P., Molyneaux, B.J., Chen, J., Inoue, J., Kominami, R., Macklis, J.D. (2005). Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron*. **45**(2): 1523-1537.
- Assenov, Y., Müller, F., Lutsik, F., Walter, J., Lengauer, T., Bock, C. (2014). Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods*. **11**(11):1138-1140.
- Bachman, M., Uribe-Lewis, S., Yang, X., Burgess, H.E., Iurlaro, M., Reik, W., Murrell, A., Balasubramanian, S. (2015). 5-Formylcytosine can be a stable DNA modification in mammals. *Nat. Chem. Biol.* **11**(8): 555–557.
- Bacquet, C., Imamura, T., Gonzalez, C.A., Conejeros, I., Kausel, G., Neildez-Nguyen, T.M., Paldi, A., Gallardo, M.H. (2008). Epigenetic processes in a tetraploid mammal. *Mamm. Genome*, **19**(6): 439–447.
- Bai, L., Yan, P., Cao, X., Jia, L., Zhang, C., Guo, D. (2015). Methylation-sensitive restriction enzyme nested real time PCR, a potential approach for sperm DNA identification. *J. Forensic Leg. Med.* **34**: 34–39.
- Barlow, D.P., Bartolomei, M.S. (2014). Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol.* **6**(2).
- Bernstein, B.E., Meissner, A., Lander, E.S. (2007). The mammalian epigenome. *Cell* **128**(4): 669–681.
- Bhogal, B., Arnaudo, A., Dymkowski, A., Best, A., Davis, T.L. (2004). Methylation at mouse *Cdkn1c* is acquired during postimplantation development and functions to maintain imprinted expression. *Genomics*. **84**(6): 961–70.
- Bird, A.P. (1978). Use of restriction enzymes to study eukaryotic DNA methylation: II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J. Mol. Biol.* **118**(1): 49–60.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**(1): 6–21.

- Bird, A. (2011). The dinucleotide CG as a genomic signalling module. *J Mol Biol.* **409**(1): 47–53.
- Bock, C. (2012). Analysing and interpreting DNA methylation data. *Nature Reviews Genetics.* **13**(10): 705-719.
- Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W., Balasubramanian S. (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science.* **336** (6083): 934–937.
- Booth, M.J., Marsico, G., Bachman, M., Beraldi, D., Balasubramanian, S. (2014). Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nat. Chem.* **6**(5): 435–440.
- Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science.* **317**(5845): 1760–1764.
- Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science.* **294**(5551): 2536-2539.
- Boyle, P., Clement, K., Gu, H., Smith, Z.D., Ziller, M., Fostel, J.L., Holmes, L., Meldrim, J., Kelley, F., Gnirke, A., Meissner, A. (2012). Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol.* **13**(10):R92.
- Brambilla, R., Gnesutta, N., Minichiello, L., White, G., Roylance, A.J., Herron, C.E., Ramsey, M., Wolfer, D.P., Cestari, V., Rossi-Arnaud, C., Grant, S.G., Chapman, P.F., Lipp, H.P., Sturani, E., Klein, R. (1997). A role for the Ras signalling pathway in synaptic transmission and long-term memory. *Nature.* **390**(6657):281-6.
- Buraschi, S., Neill, T., Goyal, A., Poluzzi, C., Smythies, J., Owens, R.T., Schaefer, L., Torres, A., Iozzo, R.V. (2013). Decorin causes autophagy in endothelial cells via Peg3. *Proc Natl Acad Sci USA.* **110**(28):E2582-E2591.
- Butler, M.G. (2009). Genomic imprinting disorders in humans: a mini-review. *J. Assist. Reprod. Genet.* **26**(9-10): 477-486.
- Cardoso, M.C., Leonhardt, H. (1999). DNA methyltransferase is actively retained in the cytoplasm during early development. *J. Cell Biol.* **147**(1): 25–32.

- Cedar, H., Solage, A., Glaser, G., Razin, A. (1979). Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI. *Nucleic Acids Res.* **6**(6): 2125–2132.
- Cen, H., Papageorge, A.G., Vass, W.C., Zhang, K.E., Lowy, D.R. (1993). Regulated and constitutive activity by CDC25Mm (GRF), aRas-specific exchange factor. *Mol Cell Biol.* **13**(12): 7718–7724.
- Chen, Z.J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* **58**: 377–406.
- Chen, Z.X., Riggs, A.D. (2011). DNA methylation and demethylation in mammals. *J Biol Chem.* **286**(21): 18347–18353.
- Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol.* **23**(16): 5594–5605.
- Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nat Rev Genet.* **6**(11):836–46.
- Davis, T.L., Yang, G.J., McCarrey, J.R., Bartolomei, M.S. (2000). The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Hum. Mol. Genet.* **9**(19): 2885–2894.
- Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.W., Page, D.C., Jaenisch, R. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Stem Cell.* **9**(2): 166–175.
- Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M., Faull, K.F., Lyko, F., Jaenisch, R. (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell.* **24**(3): 310–323.
- Deaton, A.M., Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* **25**(10): 1010–1022.

Dong, E., Guidotti, A., Grayson, D.R., Costa, E. (2007). Histone hyperacetylation induces demethylation of reelin and 67-kDa glutamic acid decarboxylase promoters. *Proc. Natl. Acad. Sci. USA*. **104**(11): 4676–4681.

Drake, N.M., Park, Y.J., Shirali, A.S., Cleland, T.A., Soloway, P.D. (2009). Imprint switch mutations at *Rasgrf1* support conflict hypothesis of imprinting and define a growth control mechanism upstream of IGF1. *Mamm Genome*. **20**(9-10): 654–663.

Duncan, A.W. (2013). Aneuploidy, polyploidy and ploidy reversal in the liver. *Semin Cell Dev Biol*. **24**(4): 347-356.

Easwaran, H.P., Schermelleh, L., Leonhardt, H., Cardoso, M.C. (2004). Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO Rep*. **5**(12):1181-6.

Edwards, C.A., Ferguson-Smith, A.C. (2007). Mechanisms regulating imprinted genes in clusters. *Curr Opin Cell Biol*. **19**(3): 281- 289.

Farnsworth, C.L., Freshney, N.W., Rosen, L.B., Ghosh, A., Green-berg, M.E., Feig, L.A. (1995). Calcium activation of Ras mediated by neuronal exchange factor Ras-GRF. *Nature*. **376**(6540): 524–527.

Feinberg, A.P., Cui, H., Ohlsson, R. (2002). DNA methylation and genomic imprinting: insights from cancer into epigenetic mechanisms. *Semin Cancer Biol*. **12**(5):389-98.

Feng, S., Jacobsen S.E., Reik W. (2010). Epigenetic reprogramming in plant and animal development. *Science*. **330**(6004): 622–627.

Feng, W., Marquez, R.T., Lu, Z., Liu, J., Lu, K.H., Issa, J.P., Fishman, D.M., Yu, Y., Bast, R.C. Jr. (2008). Imprinted tumor suppressor genes ARHI and PEG3 are the most frequently down-regulated in human ovarian cancers by loss of heterozygosity and promoter methylation. *Cancer*. **112**(7): 1489-1502.

Ferguson-Smith, A.C., Sasaki, H., Cattanach, B.M., Surani, M.A. (1993). Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature*. **362**: 751–755.

Ferrón, S.R., Charalambous, M., Radford, E., McEwen, K., Wildner, H., Hind, E., Morante-Redolat, J.M., Laborda, J., Guillemot, F., Bauer, S.R., Fariñas, I., Ferguson-

Smith, A.C. (2011) Postnatal loss of Dlk1 imprinting in stem cells and niche astrocytes regulates neurogenesis. *Nature*. **475**(7356):381-5.

Font de Mora, J., Esteban, L.M., Burks, D.J., Nunez, A., Garces, C., Garcia-Barrado, M.J., Iglesias-Osma, M.C., Moratinos, J., Ward, J.M., Santos, E. (2003). Ras-GRF1 signaling is required for normal beta-cell development and glucose homeostasis. *Embo J*. **22**(12): 3039–3049.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA*. **89**(5): 1827-1831.

Gabriel, J.M., Gray, T.A., Stubbs, L., Saitoh, S., Ohta, T., Nicholls, R.D. (1998). Structure and function correlations at the imprinted mouse *Snrpn* locus. *Mamm. Genome*. **9**(10): 788–793.

Gallardo, M.H., Bickham, J.W., Honeycutt, R.L., Ojeda, R.A., Köhler, N. (1999). Discovery of tetraploidy in a mammal. *Nature*. **401**(6751):341.

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., Chedin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell*. **45**(6): 814–825.

Ginsburg, M., Snow, M.H., McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. *Development*. **110**(2): 521–528.

Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., Brückl, T., Biel, M., Carell, T. (2010) Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One*. **5**(12):e15367.

Goll, M.G., Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem*. **74**:481–514.

Gu, T.P., Guo, F., Yang, H., Wu, H.P., Xu, G.F., Liu, W., Xie, Z.G., Shi, L., He, X., Jin, S.G., Iqbal, K., Shi, Y.G., Deng, Z., Szabó, P.E., Pfeifer, G.P., Li, J., Xu, G.L. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*. **477**(7366): 606–610.

- Guntrum, M., Vlasova, E., Davis, T.L (2017). Asymmetric DNA methylation of CpG dyads is a feature of secondary DMRs associated with the Dlk1/Gtl2 imprinting cluster in mouse. *Epigenetics & Chromatin*. **10**: 31.
- Guo, J.U., Su, Y., Shin, J.H., Shin, J., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G., Zhu, H., Chang, Q., Gao, Y., Ming, G.L., Song, H. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci*. **17**(2): 215-22.
- Hackett, J.A., Zyllicz, J.J., Surani, M.A. (2012). Parallel mechanisms of epigenetic reprogramming in the germline. *Trends Genet*. **28**(4):164–174.
- Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev*. **117**(1-2): 15–23.
- Hashimoto, K., Kokubun, S., Itoi, E., Roach, H.I. (2007). Improved quantification of DNA methylation using methylation-sensitive restriction enzymes and real-time PCR. *Epigenetics*. **2**(2): 86-91.
- Hashimoto, H., Pais, J.E., Dai, N., Corrêa, I.R. Jr., Zhang, X., Zheng, Y., Cheng, X. (2015). Structure of Naegleria Tet-like dioxygenase (NgTet1) in complexes with a reaction intermediate 5-hydroxymethylcytosine DNA. *Nucleic Acids Res*. **43**(22):10713-21.
- Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*. **405**(6785):486-489.
- Hata, K., Okano, M., Lei, H., Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development*. **129**(8): 1983-1993.
- He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C.X., Zhang, K., He, C., Xu, G.L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. **333**(6047):1303-1307.

- Herman, C.J., Lapham, L.W. (1969). Neuronal polyploidy and nuclear volumes in the cat central nervous system. *Brain Res* **15**:35-48.
- Hermann, A., Goyal, R., Jeltsch, A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J. Biol. Chem.* **279** (46): 48350-48359.
- Herzenberg, L.A., Tung, J., Moore, W.A., Herzenberg, L.A., Parks, D.R. (2006). Interpreting flow cytometry data: a guide for the perplexed. *Nat Immunol.* **7**:681–685.
- Holliday, R., Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science.* **187**(4173):226-232.
- Holoch, D., Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* **16**(2): 71–84.
- Hon, G.C., Rajagopal, N., Shen, Y., McCleary, D.F., Yue, F., Dang, M.D., Ren, B. (2013). Adult tissue methylomes harbor epigenetic memory at embryonic enhancers. *Nat Genet.* **45**(10): 1198–1206.
- Howlett, S.K., Reik, W. (1991). Methylation levels of maternal and paternal genomes during preimplantation development. *Development.* **113**(1):119-27.
- Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W., Liu, M., Sun, C., Yin, X., Li, J., Tan, X., Wang, P., Wang, Y., Fang, D., Cui, Q., Yang, P., He, C., Jiang, H., Luo, C., Xu, Y. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature.* **527**(7576):118-22.
- Hua, D., Hu, Y., Wu, Y.Y., Cheng, Z.H., Yu, J., Du, X., Huang, Z.H. (2011). Quantitative methylation analysis of multiple genes using methylation-sensitive restriction enzyme-based quantitative PCR for the detection of hepatocellular carcinoma. *Exp. Mol. Pathol.* **91**(1): 455–460.
- Huang, J.M., Kim, J. (2009) DNA methylation analysis of the mammalian PEG3 imprinted domain. *Gene.* **442**(1-2): 18-25.
- Innocenti, M., Zippel, R., Brambilla, R., Sturani, E. (1999) CDC25(Mm)/Ras-GRF1 regulates both Ras and Rac signaling pathways. *FEBS Lett.* **460**(2): 357–362.

Inoue, A., Shen, L., Dai, Q., He, C., Zhang, Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res.* 21(12): 1670–1676.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. **333**(6047): 1300–1303.

Ivanova, E., Kelsey, G. (2011). Imprinted genes and hypothalamic function. *J Mol Endocrinol.* **47**: R67-R74.

Jaenisch, R., Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33**: 245–254.

Jin, S.G., Kadam, S., Pfeifer, G.P. (2010). Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res.* **38**(11):e125.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Rev. Genet.* **13**(7):484-492.

Kagami, M., O’Sullivan, M.J., Green, A.J., Watabe, Y., Arisaka, O., Masawa, N., Matsuoka, K., Fukami, M., Matsubara, K., Kato, F., Ferguson-Smith, A.C., Ogata, T. (2010). The IG-DMR and the MEG3-DMR at human chromosome 14q32.2: hierarchical interaction and distinct functional properties as imprinting control centers. *PLoS Genet.* **6**(6):e1000992.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J.* **32**(3): 340-353.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature*. **429**(6994): 900-903.

Kaufman, M.H. (Ed.). (1992). *The Atlas of Mouse Development* (San Diego, California: Academic Press), pp. 389–407.

- Kelsey, G.; Feil, R. (2013). New insights into establishment and maintenance of DNA methylation imprints in mammals. *Phil. Trans. R. Soc. B.* **368**(1609): 20110336.
- Kohli, R.M., Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature.* **502**(7472):472-479.
- Kim, J., Ekram, M.B., Kim, H., Faisal, M., Frey, W.D., Huang, J.M., Tran, K., Kim, M.M., Yu, S. (2012). Imprinting control region (ICR) of the Peg3 domain. *Hum Mol Genet.* **21**(12):2677-2687.
- Kim, J., Frey, W.D., He, H., Kim, H., Ekram, M.B., Bakshi, A., Faisal, M., Perera, B.P.U., Ye, A., Teruyama R. (2013). Peg3 mutational effects on reproduction and placenta-specific gene families. *PLoS One.* **8**(12): e83359.
- Kiyono, M., Kaziro, Y., Satoh, T. (2000). Induction of Rac-guanine nucleotide exchange activity of Ras-GRF1/CDC25(Mm) following phosphorylation by the nonreceptor tyrosine kinase Src. *J Biol Chem.* **275**: 5441–5446.
- Kobayashi, H., Sakurai, T., Sato, S., Nakabayashi, K., Hata, K., Kono, T. (2012). Imprinted DNA methylation reprogramming during early mouse embryogenesis at the Gpr1-Zdbf2 locus is linked to long cis-intergenic transcription. *FEBS Lett.* **586**(6):827–833.
- Koziol, M.J., Bradshaw, C.R., Allen, G.E., Costa, A.S., Frezza, C., Gurdon, J.B. (2016). Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications. *Nat. Struct. Mol. Biol.* **23**(1):24-30.
- Kriaucionis, S., Heintz, N. (2009).The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science.* **324**(5929):929-30.
- Kulis, M., Merkel, A., Heath, S., Queirós, A.C., Schuyler, R.P., Castellano, G., Beekman, R., Raineri, E., Esteve, A., Clot, G., Verdaguer-Dot, N., Duran-Ferrer, M., Russiñol, N., Vilarrasa-Blasi, R., Ecker, S., Pancaldi, V., Rico, D., Agueda, L., Blanc, J., Richardson, D., Clarke, L., Datta, A., Pascual, M., Agirre, X., Prosper, F., Alignani, D., Paiva, B., Caron, G., Fest, T., Muench, M.O., Fomin, M.E., Lee, S.T., Wiemels, J.L., Valencia, A., Gut, M., Flicek, P., Stunnenberg, H.G., Siebert, R., Küppers, R., Gut, I.G., Campo, E., Martín-Subero, J.I. (2015). Whole-genome fingerprint of the DNA methylome during human B-cell differentiation. *Nat Genet.* **47**(7): 746–756.

Kumaki, Y., Oda, M., Okano, M. (2008). QUMA: quantification tool for methylation analysis. *Nucleic Acids Res.* **36**: W170-W175.

Kuo, K.C., McCune, R.A., Gehrke, C.W., Midgett, R., Ehrlich, M. (1980). Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. *Nucleic Acids Res.* **8**(20):4763–4776.

Kuroiwa, Y., Kaneko-Ishino, T., Kagitani, F., Kohda, T., Li, L.L., Tada, M., Suzuki, R., Yokoyama, M., Shiroishi, T., Wakana, S., Barton, S.C., Ishino, F., Surani, M.A. (1996). Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger protein. *Nat Genet.* **12**(2):186-190.

Laird, C.D., Pleasant, N.D., Clark, A.D., Sneed, J.L., Hassan, K.M., Manley, N.C., Vary, J.C. Jr., Morgan, T., Hansen, R.S., Stöger, R. (2004). Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc. Natl. Acad. Sci. USA.* **101**(1): 204-209.

Lan, X., Adams, C., Landers, M., Dudas, M., Krissinger, D., Marnellos, G., Bonneville, R., Xu, M., Wang, J., Huang T.H.M., Meredith, G., Jin, V.X. (2011). High resolution detection and analysis of CpG dinucleotides methylation using MBD-seq technology. *PLoS ONE.* **6**(7): e22226.

Lapham, L.W. (1968). Tetraploid DNA content of Purkinje neurons of human cerebellar cortex. *Science.* **159**: 310-312.

Leggatt, R.A., Iwama, G.K. (2003). Occurrence of polyploidy in the fishes. *Reviews in Fish Biology and Fisheries* **13**: 237-246.

Leonhardt, H., Page, A.W., Weier, H.U., Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell.* **71**(5): 865–873.

Li, E., Beard, C., Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature.* **366** (6453): 362–365.

Li, L., Keverne, E.B., Aparicio, S.A., Ishino, F., Barton, S.C., Surani, M.A. (1999). Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science.* **284**(5412):330-333.

- Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., Ferguson-Smith, A.C. (2008). A maternal-zygotic effect gene, *Zfp57*, maintains both maternal and paternal imprints. *Dev. Cell.* **15**(4): 547-557.
- Li, J.Y., Lees-Murdock, D.J., Xu, G.L., Walsh, C.P. (2004). Timing of establishment of paternal methylation imprints in the mouse. *Genomics.* **84**(6):952-960.
- Lin, S.P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J., Ferguson-Smith, A.C. (2003). Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the *Dlk1-Gtl2* imprinted cluster on mouse chromosome 12. *Nat Genet.* **35**(1): 97-102.
- Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A.H., Thomson, J.A., Ren, B., Ecker, J.R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences *Nature.* **462**(7271):315-22.
- Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., Yu, M., Tonti-Filippini, J., Heyn, H., Hu, S., Wu, J.C., Rao, A., Esteller, M., He, C., Haghghi, F.G., Sejnowski, T.J., Behrens, M.M., Ecker, J.R. (2013) Global epigenomic reconfiguration during mammalian brain development. *Science.* **341**(6146):1237905.
- Liu, Z., Wu, J., Xie, Z., Liu, S., Fan-Havard, P., Huang, T.H., Plass, C., Marcucci, G., Chan, K.K. (2009). Quantification of regional DNA methylation by liquid chromatography/tandem mass spectrometry. *Anal. Biochem.* **391**(2): 106-113.
- Lopes, S., Lewis, A., Hajkova, P., Dean, W., Oswald, J., Forné, T., Murrell, A., Constância, M., Bartolomei, M., Walter, J., Reik, W. (2003). Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Hum Mol Genet.* **12**(3):295-305.
- López-Sánchez, N., Frade, J.M. (2013). Genetic Evidence for p75^{NTR}-Dependent Tetraploidy in Cortical Projection Neurons from Adult Mice. *J Neurosci.* **33**(17): 7488-7500.
- López-Sánchez, N., Ovejero-Benito, M.C., Borreguero, L., Frade, J.M. (2011). Control of neuronal ploidy during vertebrate development. *Results Probl Cell Differ.* **53**:547-563.

López-Sánchez, N., Fontán-Lozano, Á., Pallé, A., González-Álvarez, V., Rábano, A., Trejo, J.L., Frade, J.M. (2017). Neuronal tetraploidization in the cerebral cortex correlates with reduced cognition in mice and precedes and recapitulates Alzheimer's-associated neuropathology. *Neurobiol Aging*. **56**:50-66.

López-Sánchez, N., Patiño-Parrado, I., Frade, J.M. (2017b). Quantification, isolation, and subsequent epigenetic analysis of tetraploid neurons by flow cytometry. In: *Genomic Mosaicism in neurons and other cell types* (Eds. Jose M. Frade and Fred H. Gage). Springer Verlag, Heidelberg (in press).

Lu, X., Song, C.X., Szulwach, K., Wang, Z., Weidenbacher, P., Jin, P., He, C. (2013). Chemical modification-assisted bisulfite sequencing (CAB-seq) for 5-Carboxylcytosine detection in DNA. *J.Am. Chem. Soc.* **135**(25): 9315–9317.

Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H., Trasler, J.M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. *Genomics*. **79**(4): 530-538.

Lucifero, D., Mann, M.R., Bartolomei, M.S., Trasler, J.M. (2004). Gene-specific timing and epigenetic memory in oocyte imprinting. *Hum Mol Genet.* **13**(8):839-849.

MacDonald, J.I., Verdi, J.M., Meakin, S.O. (1999). Activity-dependent interaction of the intracellular domain of rat TrkA with intermediate filament proteins, the beta-6 proteasomal subunit, Ras-GRF1, and the p162 subunit of eIF3. *J Mol Neurosci.* **13**: 141–158.

Madlung, A., Wendel, J.F. (2013). Genetic and epigenetic aspects of polyploid evolution in plants. *Cytogenet Genome Res.* **140**(2-4):270-285.

Maiti, A., Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem.* **286**(41): 35334-35338.

Marco-Sola, S., Sammeth, M., Guigó, R., Ribeca, P. (2012) The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat Methods.* **9**(12): 1185-1188.

Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., Sun, Y.E. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science.* **302**(5646):890-893.

- Mattingly, R.R., Macara, I.G. (1996). Phosphorylation-dependent activation of the Ras-GRF/CDC25Mm exchange factor by muscarinic receptors and G-protein beta gamma subunits. *Nature*. **382**: 268–272.
- Mattingly, R.R., Saini, V., Macara, I.G. (1999). Activation of the Ras-GRF/CDC25Mm exchange factor by lysophosphatidic acid. *Cell Signal*. **11**(8): 603–610.
- Meissner, A., Gnirke, A., Bell, G.W., Ramsahoye, B., Lander, E.S., Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res*. **33**(18):5868–5877.
- Morillo, S.M., Escoll, P., de la Hera, A., Frade, J.M. (2010). Somatic tetraploidy in specific chick retinal ganglion cells induced by nerve growth factor. *Proc Natl Acad Sci USA*. **107**(1):109-14.
- Morison, I.M., Ramsay, J.P., Spencer, H.G. (2005). A census of mammalian imprinting. *Trends Genet*. **21**(8):457-465.
- Mosch, B., Morawski, M., Mittag, A., Lenz, D., Tarnok, A., Arendt, T. (2007) Aneuploidy and DNA replication in the normal human brain and Alzheimer's disease. *J Neurosci*. **27**(26):6859-6867.
- Museridze, D.P., Svanidze, I.K., Macharashvili, D.N. (1975). Content of DNA and dry weight of the nuclei of neurons of the external geniculate body and retina of the eye in guinea pigs. *Sov J Dev Biol*. **5**:269-72.
- Neri, F., Incarnato, D., Krepelova, A., Rapelli, S., Anselmi, F., Parlato, C., Medana, C., Bello, F., Oliviero, S. (2015) Single-base resolution analysis of 5-formyl and 5-carboxyl cytosine reveals promoter DNA methylation dynamics. *Cell Reports*. **10**(5): 674–683.
- Nestor, C.E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J.M., Harrison, D.J., Meehan, R.R. (2012). Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res*. **22**(3): 467–477.
- Ngo, T.T., Yoo, J., Dai, Q., Zhang, Q., He, C., Aksimentiev, A., Ha, T. (2016). Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability. *Nat. Commun*. **7**: 10813.

- Nicholls, R.D., Knepper, J.L. (2001). Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. *Annu. Rev. Genomics Hum. Genet.* **2**: 153–175.
- Nunez, R. (2001). DNA measurement and cell cycle analysis by flow cytometry. *Curr Issues Mol Biol.* **3**(3):67-70.
- Ogawa, H., Wu, Q., Komiyama, J., Obata, Y., Kono, T. (2006). Disruption of parental-specific expression of imprinted genes in uniparental fetuses. *FEBS Lett.* **580**: 5377–5384.
- Okano, M., Bell, D.W., Haber, D.A., Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell.* **99**(3): 247-257.
- Olek, A., Walter, J. (1997). The pre-implantation ontogeny of the H19 methylation imprint. *Nat. Genet.* **17**: 275–276.
- Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., Cheng, X., Bestor, T.H. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature.* **448**(7154):714-717.
- Patiño-Parrado, I., Gómez-Jiménez, Á., López-Sánchez, N., Frade, J.M. (2017). Strand-specific CpG hemimethylation, a novel epigenetic modification functional for genomic imprinting. *Nucleic Acids Res.* **45**(15):8822-8834.
- Perera, B.P., Kim, J. (2016). Sex and tissue specificity of Peg3 promoters. *PLoS One.* **11**(10): e0164158.
- Plass, C., Shibata, H., Kalcheva, I., Mullins, L., Kotelevtseva, N., Mullins, J., Kato, R., Sasaki, H., Hirotsune, S., Okazaki, Y., Held, W.A., Hayashizaki, Y., Chapman, V.M. (1996) Identification of Grf1 on mouse chromosome 9 as an imprinted gene by RLGS-M. *Nat Genet.* **14**: 106–109.
- Plongthongkum, N., Diep, D.H., Zhang, K. (2014) Advances in the profiling of DNA modifications: cytosine methylation and beyond. *Nat. Rev. Genet.* **15**(10):647-661.
- Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., Reik, W. (2010). Genome wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature.* **463**(7284):1101-1105.

- Quinlan, A.R., Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26**(6):841-842.
- Quivoron, C., Couronné, L., Della Valle, V., Lopez, C.K., Plo, I., Wagner-Ballon, O., Do Cruzeiro, M., Delhommeau, F., Arnulf, B., Stern, M.H., Godley, L., Opolon, P., Tilly, H., Solary, E., Duffourd, Y., Dessen, P., Merle-Beral, H., Nguyen-Khac, F., Fontenay, M., Vainchenker, W., Bastard, C., Mercher, T., Bernard, O.A. (2011). TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell*. **20**(1):25-38.
- Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A., Trono, D. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell*. **44**(3):361-372.
- Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci USA*. **97**(10):5237–5242.
- Raiber, E.A., Murat, P., Chirgadze, D.Y., Beraldi, D., Luisi, B.F., Balasubramanian, S. (2015) 5-Formylcytosine alters the structure of the DNA double helix. *Nat Struct Mol Biol*. **22**(1): 44–49.
- Raineri, E., Dabad, M., Heath, S. (2014). A note on exact differences between beta distributions in genomic (Methylation) studies. *PLoS ONE*. **9**(5): e97349.
- Reik, W., Walter, J. (2001). Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet*. **2**(1): 21-32.
- Reik, W., Dean, W., Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science*. **293**(5532):1089-1093.
- Relaix, F., Weng, X., Marazzi, G., Yang, E., Copeland, N., Jenkins, N., Spence, S.E., Sassoon, D. (1996). Pw1, a novel zinc finger gene implicated in the myogenic and neuronal lineages. *Dev. Biol*. **177**(2):383-396.
- Riggs, A.D. (1975). X inactivation, differentiation and DNA methylation. *Cytogenet.Cell Genet*. **14**(1):9-25.

- Robinson, K.N., Manto, K., Buchsbaum, R.J., MacDonald, J.I., Meakin, S.O. (2005). Neurotrophin-dependent tyrosine phospho-rylation of Ras guanine-releasing factor 1 and associated neuriteoutgrowth is dependent on the HIKE domain of TrkA. *J Biol Chem.* **280**(1):225-35.
- Ronaghi, M., Uhlén, M., Nyrén, P. (1998). A sequencing method based on real-time pyrophosphate. *Science.* **281**(5375): 363-365.
- Saitou, M., Kagiwada, S., Kurimoto, K. (2011). Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development.* **139**(1):15–31.
- Santos, F., Hendrich, B., Reik, W., Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol.* **241**(1):172-82.
- Schermelleh, L., Haemmer, A., Spada, F., Rösing, N., Meilinger, D., Rothbauer, U., Cardoso, M.C., Leonhardt, H. (2007). Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res.* **35**(13): 4301-4312.
- Schiesser, S., Pfaffeneder, T., Sadeghian, K., Hackner, B., Steigenberger, B., Schröder, A.S., Steinbacher, J., Kashiwazaki, G., Höfner, G., Wanner, K.T., Ochsenfeld, C., Carell, T. (2013). Deamination, oxidation, and C-C bond cleavage reactivity of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine. *J Am Chem Soc.* **135**(39):14593-14599.
- Schmauss, C., Brines, M.L., Lerner, M.R. (1992) The gene encoding the small nuclear ribonucleoprotein-associated protein N is expressed at high levels in neurons. *J. Biol. Chem.*, **267**(12): 8521–8529.
- Schneider, K., Fuchs, C., Dobay, A., Rottach, A., Qin, W., Wolf, P., Álvarez-Castro, J.M., Nalaskowski, M.M., Kremmer, E., Schmid, V., Leonhardt, H., Schermelleh, L. (2013). Dissection of cell cycle-dependent dynamics of Dnmt1 by FRAP and diffusioncoupled modeling. *Nucleic Acids Res.* **41**(9): 4860–4876.
- Schultz, R., Woodfine, K., Menhenniott, T.R., Bourc'his, D., Bestor, T., Oakey, R.J. (2008). WAMIDEX: a web atlas of murine genomic imprinting and differential expression. *Epigenetics.* **3**(2):89-96.

- Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., Rauscher, F.J. 3rd. (2002). SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zincfinger proteins. *Genes Dev.* **16**(8):919-32.
- Sigl-Glöckner, J., Brecht, M. (2017). Polyploidy and the Cellular and Areal Diversity of Rat Cortical Layer 5 Pyramidal Neurons. *Cell Reports.* **20**(11): 2575–2583.
- Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., Reik, W. (2012). The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells. *Mol Cell.* **48**(6): 849–862.
- Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., Tajima, S., Mitsuya, K., Okano, M., Koseki, H. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature.* **450**(7171):908-912.
- Shemer, R., Birger, Y., Riggs, A.D., Razin, A. (1997). Structure of the imprinted mouse *Snrpn* gene and establishment of its parental-specific methylation pattern. *Proc Natl Acad Sci USA.* **94**(10): 10267–72.
- Sheng, K., Cao, W., Niu, Y., Deng, Q., Zong, C. (2017). Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat Methods.* **14**(3):267-270.
- Shirazi Fard, S., Jarrin, M., Boije, H., Fillon, V., All-Eriksson, C., Hallböök, F. (2013). Heterogenic final cell cycle by chicken retinal Lim1 horizontal progenitor cells leads to heteroploid cells with a remaining replicated genome. *PLoS One.* **8**(3): e59133.
- Shou, C., Wurmser, A., Suen, K.L., Barbacid, M., Feig, L.A., Ling, K. (1995) Differential response of the Ras exchange factor, Ras-GRF to tyrosine kinase and G protein mediated signals. *Oncogene.* **10**(10): 1887–1893.
- Solter, D. (1988). Differential imprinting and expression of maternal and paternal genomes. *Annual Review of Genetics.* **22**: 127–146.
- Smallwood, S.A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R., Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **43**(8): 811–814.

- Smallwood, S.A., Kelsey, G. (2012). De novo DNA methylation: a germ cell perspective. *Trends Genet.* **28**(1):33-42.
- Song, C.X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.H., Zhang, W., Jian, X., Wang, J., Zhang, L., Looney, T.J., Zhang, B., Godley, L.A., Hicks, L.M., Lahn, B.T., Jin, P., He, C. (2011). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotech.* **29**(1): 68–72
- Song, Q., Chen, Z.J. (2015). Epigenetic and developmental regulation in plant polyploids. *Curr. Opin. Plant Biol.* **24**: 101–109.
- Song, J., Pfeifer, G.P. (2016). Are there specific readers of oxidized 5-methylcytosine bases? *Bioessays.* **38**(10): 1038–1047.
- Sotomaru, Y., Kawase, Y., Ueda, T., Obata, Y., Suzuki, H., Domeki, I., Hatada, I., Kono, T. (2001). Disruption of imprinted expression of U2afbp-rs/U2af1-rs1 gene in mouse parthenogenetic fetuses. *J. Biol. Chem.* **276**(28): 26694–26698.
- Sotomaru, Y., Katsuzawa, Y., Hatada, I., Obata, Y., Sasaki, H., Kono, T. (2002). Unregulated expression of the imprinted genes H19 and Igf2r in mouse uniparental fetuses. *J. Biol. Chem.* **277**(14): 12474–12478.
- Song, J., Rechko, O., Bestor, T.H., Patel, D.J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science.* **331**(6020): 1036–1040.
- Song, C.X., Szulwach, K.E., Dai, Q., Fu, Y., Mao, S.Q., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H., Gao, J., Liu, P., Li, L., Xu, G.L., Jin, P., He, C. (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell.* **153**(3): 678-691.
- Song, C.X., He, C. (2013). Potential functional roles of DNA demethylation intermediates. *Trends Biochem Sci.* **38**(10): 480-484.
- Spahn, L., Barlow, D.P. (2003). An ICE pattern crystallizes. *Nat Genet.* **35**:11-12.
- Stöger, R.; Kubicka, P.; Liu, C.-G.; Kafri, T.; Razin, A.; Cedar, H.; Barlow, D.P. (1993). Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the expressed locus as carrying the imprinting signal. *Cell.* **73**(1): 61–71.

- Sturani, E., Abbondio, A., Branduardi, P., Ferrari, C., Zippel, R., Martegani, E., Vanoni, M., Denis-Donini, S. (1997). The Ras Guanine nucleotide exchange factor CDC25Mm is present at the synaptic junction. *Exp Cell Res.* **235**: 117–123.
- Swaney, W.T., Curley, J.P., Champagne, F.A., Keverne, E.B. (2008). The paternally expressed gene Peg3 regulates sexual experience-dependent preferences for estrous odors. *Behav Neurosci.* **122**(5): 963-973.
- Swaney, W.T., Curley, J.P., Champagne, F.A., Keverne, E.B. (2007). Genomic imprinting mediates sexual experience-dependent olfactory learning in male mice. *Proc Natl Acad Sci USA.* **104**: 6084- 6089.
- Swartz, F.J., Bhatnagar, K.P. (1981). Are CNS neurons polyploid? A critical analysis based upon cytophotometric study of the DNA content of cerebellar and olfactory bulbar neurons of the bat. *Brain Res.* **208**(2): 267- 81.
- Szwagierczak, A., Bultmann, S., Schmidt, C.S., Spada, F., Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **38**(19): e181.
- Szulwach, K.E., Li, X., Li, Y., Song, Q.X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., Vasanthakumar, A., Godley, L.A., Chang, Q., Cheng, X., He, C., Jin, P. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nature Neurosci.* **14**: 1607–1616.
- Takada, S., Paulsen, M., Tevendale, M., Tsai, C.E., Kelsey, G., Cattanach, B.M., Ferguson-Smith, A.C. (2002). Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2-H19. *Hum Mol Genet.* **11**(1): 77–86.
- Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science.* **324** (5929): 930–935.
- Thiaville, M.M., Huang, J.M., Kim, H., Ekram, M.B., Roh, T.Y., Kim, J. (2013). DNA-binding motif and target genes of the imprinted transcription factor PEG3. *Gene.* **512**(2): 314-320.

Tremblay, K.D., Duran, K.L., Bartolomei, M.S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Mol Cell Biol.* **17**:4322–4329.

Ullah, Z., Kohn, M.J., Yagi, R., Vassilev, L.T., DePamphilis, M.L. (2008). Differentiation of trophoblast stem cells into giant cells is triggered by p57/Kip2 inhibition of CDK1 activity. *Genes Dev.* **22**(21): 3024-3036.

Ullah, Z., Lee, C.Y., Depamphilis, M.L. (2009). Cip/Kip cyclin-dependent protein kinase inhibitors and the road to polyploidy. *Cell Div.* **4**:10.

Valinluck, V., Tsai, H.H., Rogstad, D.K., Burdzy, A., Bird, A., Sowers, L.C. (2004). Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res* **32**(14): 4100–4108.

Valinluck, V., Sowers, L.C. (2007). Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res.* **67**(3): 946–950.

Vernon, J.A., Butsch, J. (1957). Effect of tetraploidy on learning and retention in the salamander. *Science.* **125**(3256): 1033-1034.

Vermeulen, K., Berneman Zwi, N., Van Bockstaele, D.R. (2003). Cell cycle and apoptosis *Cell Prolif.* **36**(3): 165-175.

Vitrat, N., Cohen-Solal, K., Pique, C., Le Couedic, J.P., Norol, F., Larsen, A.K., Katz, A., Vainchenker, W., Debili, N. (1998). Endomitosis of human megakaryocytes are due to abortive mitosis. *Blood.* **91**(10): 3711-23.

Wang, J., Xia, Y., Li, L., Gong, D., Yao, Y., Luo, H., Lu, H., Yi, N. , Wu, H., Zhang, X., Tao, Q., Gao, F. (2013). Double restriction-enzyme digestion improves the coverage and accuracy of genome-wide CpG methylation profiling by reduced representation bisulfite sequencing. *BMC Genomics.* **14**: 11.

Warnecke, P.M., Biniszkievicz, D., Jaenisch, R., Frommer, M., Clark, S.J. (1998). Sequence-specific methylation of the mouse H19 gene in embryonic cells deficient in the Dnmt-1 gene. *Dev. Genet.* **22**(2): 111–121.

- Wigler, M., Levy, D., Perucho, M. (1981). The somatic replication of DNA methylation. *Cell*. **24**(1):33–40.
- Wu, H., Xu, T., Feng, H., Chen, L., Li, B., Yao, B., Qin, Z., Jin, P., Conneely, K.N. (2015). Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates *Nucleic Acids Research*. **43**(21): e141.
- Wu, S.C., Zhang, Y. (2010). Active DNA demethylation: Many roads lead to Rome. *Nat Rev Mol Cell Biol*. **11**(9): 607–620.
- Xie, W., Barr, C.L., Kim, A., Yue, F., Lee, A.Y., Eubanks, J., Dempster, E.L., Ren, B. (2012). Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell*. **148**(4): 816-31.
- Xiong, Z., Laird, P.W. (1997). COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res*. **25**(12): 2532-2534.
- Yamaguchi, A., Taniguchi, M., Hori, O., Ogawa, S., Tojo, N., Matsuoka, N., Miyake, S., Kasai, K., Sugimoto, H., Tamatani, M., Yamashita, T., Tohyama, M. (2002). Peg3/Pw1 is involved in p53-mediated cell death pathway in brain ischemia/hypoxia. *J Biol Chem*. **277**:623-629.
- Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D., Zhang, K., Zhang, Y. (2012). Tet1 controls meiosis by regulating meiotic gene expression. *Nature*. **492**(7429): 443–447.
- Yamaguchi, S., Shen, L., Liu, Y., Sendler, D., Zhang, Y. (2013). Role of Tet1 in erasure of genomic imprinting. *Nature*. **504**(7480): 460-464.
- Yamazaki, W., Amano, T., Bai, H., Takahashi, M., Kawahara, M. (2016). The Influence of Polyploidy and Genome Composition on Genomic Imprinting in Mice. *J Biol Chem*. **291**(40):20924-20931.
- Yoon, B.J., Herman, H., Sikora, A., Smith, L.T., Plass, C., Soloway, P.D. (2002). Regulation of DNA methylation of Rasgrf1. *Nat Genet*. **30**(1): 92–96.
- Yoon, B., Herman, H., Hu, B., Park, Y.J., Lindroth, A., Bell, A., West, A.G., Chang, Y., Stablewski, A., Piel, J.C., Loukinov, D.I., Lobanenko, V.V., Soloway, P.D. (2005).

Rasgrf1 imprinting is regulated by a CTCF-dependent methylation-sensitive enhancer blocker. *Mol Cell Biol.* **25**(24):11184–11190.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Jin, P., Ren, B., He, C. (2012). Tet-assisted bisulfite sequencing 5-hydroxymethylcytosine. *Nature Protoc.* **7**(12): 2159–2170.

Zanet, J., Freije, A., Ruiz, M., Coulon, V., Sanz, J.R., Chiesa, J., Gandarillas, A. (2010). A mitosis block links active cell cycle with human epidermal differentiation and results in endoreplication. *PLoS One.* **5**(12):e15701.

Zhao, M.T., Whyte, J.J., Hopkins, G.M., Kirk, M.D., Prather, R.S. (2014) Methylated DNA immunoprecipitation and high-throughput sequencing (MeDIP-seq) using low amounts of genomic DNA. *Cell Reprogram.* **16**(3):175-184.

Ziller, M.J., Muller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C.B., Bernstein, B.E., Lengauer, T., Gnirke, A, Meissner, A. (2011). Genomic Distribution and Inter-Sample Variation of Non-CpG Methylation across Human Cell Types. *PLoS Genet.* **7**(12):e1002389.

Zippel, R., Balestrini, M., Lomazzi, M., Sturani, E. (2000). Calcium and calmodulin are essential for Ras-GRF1-mediated activation of the Ras pathway by lysophosphatidic acid. *Exp Cell Res.* **258**(2):403-8.

8. Anexos

Publicaciones:

- Patiño-Parrado, I.; Gómez-Jiménez, A.; López-Sánchez, N.; Frade, J.M. (2017). Strand-specific CpG hemimethylation, a novel epigenetic modification functional for genomic imprinting. *Nucleic Acid Research*. **45** (15): 8822–8834.
- López-Sánchez, N., Patiño-Parrado, I., Frade, J.M. (2017). Quantification, isolation, and subsequent epigenetic analysis of tetraploid neurons by flow cytometry. In: *Genomic Mosaicism in neurons and other cell types*, pp 57-80 (Eds. Jose M. Frade and Fred H. Gage). Springer Verlag, Heidelberg.

Strand-specific CpG hemimethylation, a novel epigenetic modification functional for genomic imprinting

Iris Patiño-Parrado, Álvaro Gómez-Jiménez, Noelia López-Sánchez and José M. Frade*

Department of Molecular, Cellular, and Developmental Neurobiology, Cajal Institute, Consejo Superior de Investigaciones Científicas (IC-CSIC), Madrid E-28002, Spain

Received August 01, 2016; Revised May 30, 2017; Editorial Decision May 31, 2017; Accepted June 01, 2017

ABSTRACT

Imprinted genes are regulated by allele-specific differentially DNA-methylated regions (DMRs). Epigenetic methylation of the CpGs constituting these DMRs is established in the germline, resulting in a 5-methylcytosine-specific pattern that is tightly maintained in somatic tissues. Here, we show a novel epigenetic mark, characterized by strand-specific hemimethylation of contiguous CpG sites affecting the germline DMR of the murine *Peg3*, but not *Snrpn*, imprinted domain. This modification is enriched in tetraploid cortical neurons, a cell type where evidence for a small proportion of formylmethylated CpG sites within the *Peg3*-controlling DMR is also provided. Single nucleotide polymorphism (SNP)-based transcriptional analysis indicated that these epigenetic modifications participate in the maintenance of the monoallelic expression pattern of the *Peg3* imprinted gene. Our results unexpectedly demonstrate that the methylation pattern observed in DMRs controlling defined imprinting regions can be modified in somatic cells, resulting in a novel epigenetic modification that gives rise to strand-specific hemimethylated domains functional for genomic imprinting. We anticipate the existence of a novel molecular mechanism regulating the transition from fully methylated CpGs to strand-specific hemimethylated CpGs.

INTRODUCTION

In the mouse, one hundred fifty one genes are known to be imprinted (see <http://www.mousebook.org/imprinting-gene-list>). As in other mammals, most of these imprinted genes are found in clusters distributed throughout the whole murine genome (1), and they are co-regulated by cis-acting imprinting control elements that contain allele-specific, dif-

ferentially DNA-methylated regions, hereafter referred to as DMRs. Epigenetic cytosine methylation in the CpGs constituting these DMRs specifically occurs in the germline, and these 5-methylcytosine (5mC)-modified regions are referred to as germline DMRs as opposed to somatic DMRs generated after fertilization (2). The methylation pattern acquired in the germline is tightly maintained in all somatic tissues due to the activity of DNMT1, a DNA methyltransferase that methylates the newly synthesized strand after DNA replication (3,4).

Besides 5mC (5), 5-formylcytosine (5fC) can also be found as a stable DNA modification in mammalian genomes (6). This molecule is generated by sequential oxidation of 5mC which can be converted to 5-hydroxymethylcytosine (5hmC) and then to 5fC by the TET (ten eleven translocation) proteins (7). To study whether the methylation pattern of DMRs controlling genomic imprinting can be altered and whether 5fC can be present in these DMRs we focused on the *Peg3* and *Snrpn* imprinted domains. This analysis was primarily focused on neurons from the mouse cerebral cortex as these imprinted domains contain genes that are expressed in neurons (8–10). We were particularly interested in determining whether cytosine modification in cortical neurons could be altered in response to changes in ploidy level, as a small proportion of cortical neurons contain double the normal amount of DNA in their nuclei (11).

In the vertebrate nervous system, a subset of differentiating projection neurons replicate their genome as they differentiate and then remain as tetraploid (4C) neurons in the adult brain (11–13). The mechanism leading to neuronal tetraploidy depends on the activation of the neurotrophin receptor p75 (11,12), which induces the activation of p38^{MAPK} and the subsequent phosphorylation of E2F4 (14), a key regulator of the cell cycle. After DNA replication, Cdk1 inactivation prevents differentiating 4C neurons from entering mitosis (15), while p27^{Kip1} expression in these cells avoids further rounds of endoreplication (16). Somatic 4C neurons show extensive dendritic arbors and enlarged cell bodies (12), express the immediate early genes *Erg-1* and *c-*

*To whom correspondence should be addressed. Tel: +34 91 585 4740; Fax: +34 91 585 4754; Email: frade@cajal.csic.es

Fos (11), known to respond to neuronal activity (17,18), and constitute specific neuronal populations that innervate defined target areas (12), thus indicating that 4C neurons are fully functional.

In this study, we show that the germline DMR of *Peg3*, but not *Snrpn*, imprinting domain show a specific hemimethylated pattern in subpopulations of both diploid (2C) and 4C cortical neurons, being enriched in the latter group. We conclude that, in contrast to previous belief, germline DMRs can undergo covalent modifications that give rise to a novel epigenetic mark, characterized by strand-specific CpG hemimethylation. This epigenetic modification participates in the maintenance of the monoallelic expression pattern of the *Peg3* imprinted gene.

MATERIALS AND METHODS

Mice

Male C57BL/6J mice from embryonic day 11 (E11) and postnatal day 15 (P15) were used in this study. Embryos were staged as described by Kaufman (19). In embryos, sex was determined by genomic PCR, using chromosome Y-specific primers (see Table 1). DBA/2J mice were purchased from The Jackson Laboratory (Bar Harbor, ME, USA) and bred to C57BL/6J mice to produce P15 mice of hybrid genetic background. Experimental procedures were approved by the CSIC Bioethics Committee, and they were performed in accordance with the European Union guidelines.

Primary and secondary antibodies

The mouse anti-NeuN monoclonal antibody (Millipore) was used at 1/1000 and the Alexa Fluor 488 goat anti-mouse antibody was diluted 1/500 for nuclei sorting.

Oligonucleotides

Oligonucleotides used in this study are summarized in Table 1. Their specificity was verified by BLAST searches as well as by the actual PCR reaction.

Synthesis of non-methylated, hemimethylated, fully methylated, and methylformylated double-stranded DNA (dsDNA) for HhaI cleavage assays

Non-methylated and fully methylated dsDNA fragments were generated by PCR using genomic DNA isolated from cortical neurons. To this aim, dATP/dTTP/dGTP (0.2 mM each, Biotools) plus either CTP (0.2 mM, Biotools) or 5mCTP (0.2 mM, New England Biolabs) were used together with DreamTaq DNA Polymerase (1.25 units, Thermo) as well as the HhaI forward and HhaI reverse primers (0.4 μ M each, Table 1) flanking the HhaI restriction sites from the *Peg3*-controlling DMR. Amplifications were performed using the following thermal cycle: 94°C for 3 min, 50 \times (94°C for 30 s, 59°C for 1 min, 72°C for 1 min), 72°C for 20 min. PCR products were cleaned using DCC Clean & Concentrator-5 (ZymoResearch) following manufacturer's instructions. Hemimethylated and methylformylated dsDNA fragments were generated by PCR using fully methylated dsDNA. To this aim, dATP/dTTP/dGTP (0.2

mM each, Biotools) plus either CTP (0.2 mM, Biotools) or 5fCTP (0.2 mM, TriLink) were used together with DreamTaq DNA Polymerase (1.25 units, Thermo) as well as the HhaI forward and HhaI reverse primers (0.4 μ M each) described above. Amplifications were performed using the following thermal cycle: 94°C for 3 min, 1 \times (94°C for 30 s, 59°C for 1 min, 72°C for 1 min), 72°C for 20 min. PCR products were cleaned using DCC Clean & Concentrator-5 (ZymoResearch) following manufacturer's instructions. Non-methylated, hemimethylated, fully methylated, or methylformylated PCR products (10 μ l each) were incubated for 1 h at 37°C in either the presence or absence of 20 U HhaI (Thermo Fisher Scientific). After incubation, reactions were inactivated for 20 min at 80°C and run in an agarose gel.

Synthesis of 5fC-containing DNA for sodium borohydride (BH)- and O-hydroxyethylamine (EA)-based bisulfite protection assays

DsDNA containing 5fC for BH/EA-based bisulfite protection assays was synthesized as previously described by (20). To this aim, template DNA (CTCACCCACAACCA CAAACAAATTTAATACGATTAAATAATATTAATA TATTATCGATTAGTAGGTAAAGTAAGGGTATTTG ATGTGATGGGTGGTATGG) was PCR amplified with DreamTaq DNA Polymerase (Thermo) using the 5fC forward and 5fC reverse primers (Table 1), and commercially available 5fCTP (TriLink).

Cell nuclei isolation

Cell nuclei isolation was performed as described by López-Sánchez and Frade (11). Briefly, one hemicortex from mice or mouse embryos of the specified age, previously snap-frozen on dry ice and stored at -80°C, was placed in 2 ml of ice-cold, DNase-free PBS containing 0.1% Triton X-100 (Sigma) and protease inhibitor mixture (Roche) (nuclear isolation buffer). Cell nuclei were then isolated by mechanical disaggregation using a Dounce homogenizer. Undissociated tissue was removed by centrifugation at 200 \times g for 1.5 min at 4°C. The supernatant was 3-fold diluted with nuclear isolation buffer and centrifuged at 400 \times g for 4 min at 4°C. Supernatant with cellular debris was discarded, and the pellet incubated at 4°C in 50 μ l of cold nuclear isolation buffer for at least 1 h, before mechanical disaggregation by gently swirl of the vial. The quality and purity of the isolated nuclei was analyzed microscopically after staining with 100 ng/ml DAPI. A similar procedure was used for the isolation of hepatocyte nuclei.

Cell nuclei sorting

Neuronal nuclear immunostaining was performed by adding primary and secondary antibodies to isolated unfixed nuclei containing 5% FCS and 1.25 mg/ml BSA. In control samples, the primary antibodies were excluded. Finally, the reaction was incubated O/N at 4°C in the dark. Immunostained nuclei (or hepatocyte nuclei) were filtered through a 40 μ m nylon filter, and the volume adjusted with PBS containing propidium iodide (PI; Sigma) and DNase-free RNase I (Sigma) at a final concentration of 40 and 25

Table 1. Oligonucleotides used in this study

Primer	Sequence (5' → 3')	Modification	Purpose	Annealing temp. (°C)
<i>Peg3</i> up-A	TTTTGTAGAGGATTTTGATAAGGAG	5'-biotinilated	Cloning	55
<i>Peg3</i> down-A	CAATCTAATACACCCACACTAAACC		Cloning/ Pyrosequencing	55
<i>Peg3</i> down-A-biot	CAATCTAATACACCCACACTAAACC		Pyrosequencing	55
<i>Peg3</i> seq-A	ATGTTTATTTTGGGTTGGTGG		Pyrosequencing	
<i>Peg3</i> up-B	GAGGAGAAGCGGAGAGATGT	FAM/TAMRA	Enzyme-qPCR	60
<i>Peg3</i> down-B	CACAGCACTCTACGCACACA		Enzyme-qPCR	60
<i>Peg3</i> probe	AGACTGCCGAGGTCGG		Enzyme-qPCR	
<i>Peg3</i> up-C	AGAYGTTGGGGAGTTAGGAG		Hairpin-bisulfite PCR	56
<i>Peg3</i> down-C	YAAAAAATATCCACCCTAAACTAATAAC	5'-biotinilated	Hairpin-bisulfite PCR	56
<i>Peg3</i> hairpin	GCCGAGTCTGACTTTTTTGTCTAGACT		Hairpin	
<i>Peg3</i> up-D	AATGGCACATGCCTGGAAC		SNP pyrosequencing	58
<i>Peg3</i> down-D	CGATGAGTGGCCTTGTGTCA		SNP pyrosequencing	58
<i>Peg3</i> seq-B	CTCCTGTTCACTTCTTTGAGAGAC	5'-biotinilated	SNP pyrosequencing	N/A
<i>Peg3</i> seq-C	TGAGGGTCTCACTATGTAGGTGT		SNP pyrosequencing	N/A
<i>Snrpn</i> up-A	TTAGAGGGATAGAGATTTTGTATTG		Cloning/ Pyrosequencing	56
<i>Snrpn</i> down-A	CTAAAATCCACAAACCACTAAC		Cloning	56
<i>Snrpn</i> down-A-biot	CTAAAATCCACAAACCACTAAC	5'-biotinilated	Pyrosequencing	56
<i>Snrpn</i> seq	GTATGTGTAGTTATTGTTTGGGA		Pyrosequencing	56
Chromosome Y-1	GCATTTGCCTGTCAGAGAGAG		Sex determination	58
Chromosome Y-2	ACTGCTGCTGCTTTTCCAATA		Sex determination	58
5fC forward	MTMAMCCAMAAACMAMAAAMA		5fC-containing DNA	57
5fC reverse	CCATACCAACCATCACATCA		5fC-containing DNA	57
5fC forward post-bisulfite	ACACTGACGACATGGTCTACACTCACT		NGS	57
	TACAATCACAAACA			
5fC reverse post-bisulfite	TACGGTAGCAGAGACTTGGTCTCCATAC		NGS	57
	CACCCATCACATCA			
HhaI forward	CAGAGGACCCTGACAAGGAG		HhaI restriction assay	59
HhaI reverse	AGTTCAGATGGTGTGTTGGGG		HhaI restriction assay	59
N/A: not applicable.				

μg/ml, respectively, and incubated for 30 min at RT. The quality of the nuclei and specificity of immunostaining signal was checked by fluorescence microscopy. Nuclei sorting was performed with a FACS Aria cytometer (BD Biosciences) equipped with a 488-nm Coherent Sapphire solid state and 633-nm JDS Uniphase HeNe air-cooled laser. Data were collected by using a linear digital signal process. The emission filters used were BP 530/30 for Alexa 488 and BP 616/23 for PI. Data were analyzed with FACSDiva (BD Biosciences). Gating was performed by removing doublets and selecting the NeuN-positive nuclei, as previously described (11). Cellular debris, which was clearly differentiated from nuclei due to its inability to incorporate PI, was gated and excluded from the analysis. For isolation of the neuronal population highly enriched in 4C DNA amount (~90% purity), two rounds of sorting were performed. Cell nuclei from the first round of sorting were collected in Protein LoBind tubes (Eppendorf), while those from the second round were collected in DNA LoBind tubes (Eppendorf).

Genomic DNA isolation

Genomic DNA was isolated using DNeasy Blood & Tissue Kit (Qiagen) following the protocol described by the manufacturer. Two mice were used to generate each genomic DNA and at least three different genomic DNAs were employed in each experiment.

Whole genome amplification and CGH array

DNA was amplified using the GenomePlex WGA2 kit (Sigma Aldrich) according manufacturer's instructions. The amplified DNA was clean with phenol-chloroform using a standard protocol. CGH arrays were performed using the SurePrint G3 Mouse CGH Microarray Kit 180K (Agilent Technologies, Palo Alto, CA, USA) covering the whole

genome with a 10.9 Kb overall median probe spacing following the manufacturer's protocol.

Methylation-sensitive restriction enzyme-based quantitative PCR (qPCR)

Quantification of CpG island methylation in the *Peg3* DMR was performed in DNA isolated from 2C and 4C sorted neuronal nuclei as described above, using a modification of the procedure described by Hashimoto *et al.* (21). For each DNA sample, two parallel reactions were prepared, containing 15 ng of DNA (quantified by band intensity in agarose gels) in either the presence or the absence of 20 U HhaI (Thermo Fisher Scientific). All reactions were incubated for 5 h at 37°C in 20 μl total volume. After digestion, the reactions were inactivated by incubating the mixture at 80°C for 20 min. 1.14 μl of each reaction was added to 20 μl total volume PCR reaction containing 10 μl 2X TaqMan Universal PCR Master Mix, both primers at a concentration of 900 nM and a fluorescent probe at a concentration of 250 nM (for details, see list of primers in Table 1). Samples were amplified using an ABI Prism 7000 thermal cycler (Applied Biosystems) using the following conditions: 2 min at 50°C, 10 min at 95°C, followed by 40 cycles of PCR consisting of 15 s at 95°C and 1 min at 60°C. Under these conditions, qPCR efficiency was 99.96% (amplification factor: 2). Quantification was performed within linear range (R^2 of standard curve = 0.9918). In all cases, technical triplicates were performed (2C intra-assay mean CV = 18.53%; 4C intra-assay mean CV = 22.74%). Outliers were identified with the Dixon's Q test. Negative controls (DNA replaced by water) did not show amplification.

DNA reduction with BH/EA

Genomic DNA and 5fC-containing synthetic DNA was reduced using BH as previously described by Booth *et al.* (20).

Briefly, 5 μ l of an aqueous sodium borohydride (Merk) solution (1 M), freshly prepared before every reduction, was added to 100–200 ng DNA (dissolved in 15 μ l water). The reaction was vortexed and centrifuged, and held at room temperature in the dark for 1 h. The reactions were vortexed and centrifuged every 15 min to remove the bubbles that were generated, and the pressure was released by opening the lids. The reaction was finally quenched by slowly adding 10 μ l of a solution containing 750 mM sodium acetate (Sigma), pH 5.0. Since violent release of hydrogen occurs during this quenching step, the reaction was held at room temperature for 10 min or until no further gas was released. EA protection of 5fC in genomic DNA and 5fC-containing synthetic DNA was performed with 10 mM *O*-ethylhydroxylamine (Sigma-Aldrich) in 100 mM MES buffer, pH 5.0, at 37°C for 2 h, as described by Song *et al.* (22). The DNA substrates were purified with Genomic DNA Clean & Concentrator™-10 (ZymoResearch) and subjected to sodium bisulfite treatment as described below.

Bisulfite conversion

A volume of 20 μ l DNA was converted using Epitect Bisulfite Kit (Qiagen). Bisulfite conversion was performed in 85 μ l of Bisulfite Mix solution, and 35 μ l of DNA Protect Buffer. The total reaction volume was 140 μ l. The procedure was performed according to the manufacturer's instructions. A modified protocol was used for bisulfite conversion of borohydride-treated DNA. Briefly, 100–200 ng DNA (either treated with borohydride or untreated) was made up to 30 μ l in water, and then combined with bisulfite mix (80 μ l) and DNA protect buffer (30 μ l). The reactions were then subjected to two cycles of the formalin-fixed, paraffin-embedded (FFPE) thermal cycle, and worked up as per the Qiagen Epitect Bisulfite Kit protocol for FFPE samples.

Bisulfite-converted DNA amplification, cloning, and Sanger sequencing

Primers were designed to match the DMRs of the imprinted *Snrpn* and *Peg3* regions previously described by Lucifero *et al.* (8) (for sequences and annealing temperatures, see list of primers in Table 1). These primers are specific for the bisulfite-converted sequence, and they encompass 13 CpG sites within the *Snrpn* imprinted domain [corresponding to the CpG sites number 4 to 16 from Lucifero *et al.* (8)], and 12 CpG sites along the *Peg3* imprinted domain [corresponding to the first 12 CpG sites from Lucifero *et al.* (8)]. PCR amplification was performed using standard procedures. PCR products were subcloned into the pGEM-Teasy vector (Promega), and Sanger sequenced.

Pyrosequencing analysis of bisulfite converted DNA

Bisulfite-treated DNA was PCR amplified using *Snrpn*- or *Peg3*-specific primers, one of which was biotinylated at its 5' end (for sequences and annealing temperatures, see list of primers in Table 1). The PCR product was then bound to streptavidin-coated Sepharose beads (GE-Healthcare) through the biotin tag and denatured to generate single-stranded DNA to allow annealing of an internal sequencing

primer (16 pmol of each per reaction) (see list of primers in Table 1). Pyrosequencing was performed in a PSQ™96MA instrument (Biotage) using 25 μ l of amplified DNA product and PyroMark Gold Q96 reagents (Qiagen). Pyrosequencing data analysis was done using the Pyro Q-CpG software (Biotage).

Next-generation sequencing (NGS) of bisulfite-converted DNA

DsDNA containing 5fC for BH/EA-based bisulfite protection assays was reduced with either BH or EA and subjected to bisulfite conversion as described above. PCR library amplification was performed with DNA Polymerase (Biotools) using the 5fC forward post-bisulfite and 5fC reverse post-bisulfite primers (Table 1). Illumina adapter sequences included at the 5' end of the primers were used to prepare libraries, which were then sequenced on the Illumina Miseq platform (2 × 300 bp paired-end reads). Sequences were aligned using the MiSeq Reporter software (Illumina) and BAM alignments were visualized using the Integrative Genomics Viewer platform. For methylated CpG level analysis within the *Peg3* genomic region, including the *Peg3* imprinted domain-controlling DMR (PCD), genomic DNA isolated from ~2000 4C neuronal nuclei was subjected to bisulfite treatment and amplification with the Pico Methyl-Seq Library Prep kit (Zymo Research). The bisulfite-Seq library generated was sequenced on the Illumina HiSeq 2500 platform. Read mapping, processing, and analysis were performed as described previously (23).

Hairpin-bisulfite PCR

The presence of hemimethylated CpGs in the PCD was analyzed using a modification of the protocol described by Laird *et al.* (24). Briefly, 50 ng of DNA isolated from 4C sorted neuronal nuclei was digested with 2.5 U BbvI (Thermo Fisher Scientific) for 1 h at 37°C. Then, the hairpin linker (2 μ g) was ligated to the cleaved DNA using 10 U T4 DNA ligase (Thermo Fisher Scientific). Hairpin-ligated genomic DNA was then subjected to bisulfite conversion using Epitect Bisulfite Kit (Qiagen). To avoid rapid fold-back renaturation of hairpin DNA, two extra thermal-denaturation steps were introduced without modifying the total incubation time indicated by the manufacturer (incubation conditions: 5 min at 95°C, 25 min at 60°C, 5 min at 95°C, 60 min at 60°C, 5 min at 95°C, 60 min at 60°C, 5 min at 95°C, 80 min at 60°C). Then, bisulfite-converted DNA was PCR amplified with the appropriate primers (see list of primers in Table 1) using standard conditions, with the exception that the concentration of MgCl₂ was raised to 5 mM. PCR products were subcloned in the pGEM-Teasy Vector (Promega), and clones containing the appropriate inserts were Sanger sequenced.

SNP imprinting analysis

A polymorphism for SNP imprinting analysis was identified in intron 7 of mouse *Peg3* gene using the Mouse Genome Informatics tool (The Jackson Laboratory) (<http://www.>

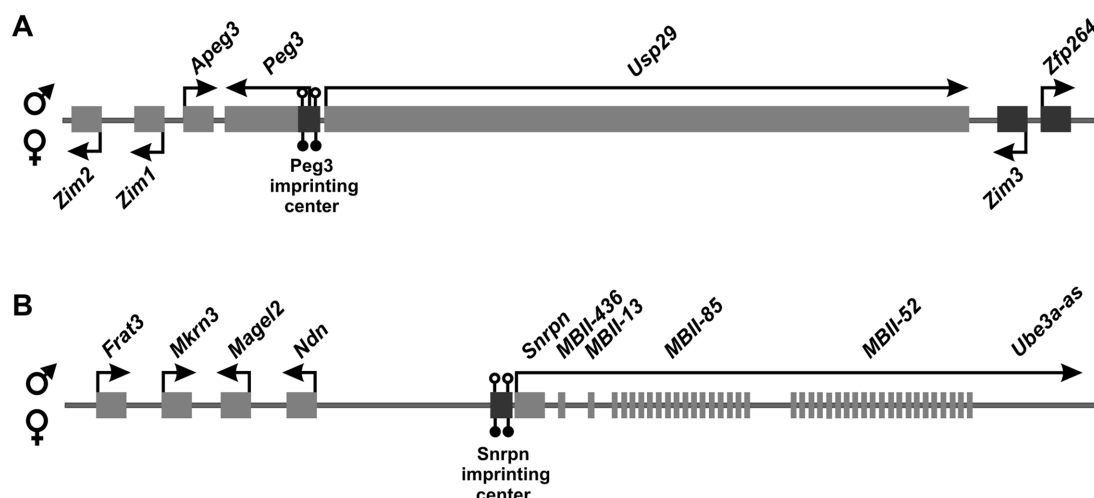


Figure 1. Scheme showing the genomic structure of murine *Peg3* and *Snrpn* imprinted domains. (A) Genes constituting the *Peg3* imprinted domain are shown. *Apeg3*, *Peg3*, *Usp29* and *Zfp264* show paternal expression whereas *Zim1*, *Zim2* and *Zim3* are expressed from the maternal-derived chromosome. The *Peg3* imprinting center, which contains the *Peg3*-controlling DMR, is located within the *Peg3* promoter region. Methylated CpG sites from the *Peg3*-controlling DMR are represented by solid circles. Open circles represent non-methylated CpGs from the *Peg3*-controlling DMR. (B) Genes constituting the *Snrpn* imprinted domain are shown. All transcriptional units show paternal expression. The *Snrpn* imprinting center, which contains the *Snrpn*-controlling DMR, is located within the *Snrpn* promoter region. Methylated CpG sites from the *Snrpn*-controlling DMR are represented by solid circles. Open circles represent non-methylated CpGs from the *Snrpn*-controlling DMR.

informatics.jax.org/strains.SNPs.shtml). This SNP (NCBI accession number rs31442449), characterized by a nucleotide change from T (in C57BL/6J background) to C (in DBA/2J background), is located at bp 18 192 of the mouse *Peg3* gene (NCBI accession number NC_000073, complement to bp 6 705 960–6 730 419). Total RNA was isolated using the RNeasy-4PCR Kit (Ambion) from 2C and 4C neuronal nuclei derived from the cerebral cortex of P15 male littermates obtained from the mating of DBA/2J and C57BL/6J mice. Cerebral cortices were treated with All-protect (Qiagen) following the protocol described by the manufacturer before nuclear isolation. RNA samples were treated with DNase I (Thermo Fisher Scientific) to remove traces of genomic DNA contamination. Synthesis of cDNA was performed using SuperScript III Reverse Transcriptase (Invitrogen) following manufacturer's instructions. PCRs were performed using *Peg3*-specific primers, one of which was biotinylated at its 5' end (for sequence and amplification conditions, see list of primers in Table 1). PCR products were pyrosequenced with an internal sequencing primer (see list of primers in Table 1) as indicated above.

Data analysis and statistics

Sanger sequencing results were analyzed for change in methylation pattern with the Quantification tool for Methylation Analysis (Quma) (<http://quma.cdb.riken.jp/>) (25). The software generates a panel of solid and open circles indicating methylated and unmethylated CpGs, respectively. Quantitative data were obtained from DNA analysis in triplicate from at least three different DNAs (obtained from two mice) for each experimental point, except in BH experiments with P20 liver, in which one DNA sample for each experimental point was used. Statistical analyses and graphics were performed using the Graphpad Prism 5 software package. Global methylation analysis was performed

by pooling data from all CpG sites in each particular DMR. qPCR data were analyzed using 7000 System Sequence Detection Software version 1.2.3 (Applied Biosystems). The percentage of methylation in methylation-sensitive restriction qPCR assays was estimated by dividing the values obtained with the digested DNA aliquots by those of the non-digested DNA aliquots. The average values were obtained from thirteen (2C) and six (4C) independent experiments in triplicate. Statistical significance was determined using two-tailed Student's *t* test unless otherwise specified.

RESULTS

Reduced methylation levels in the *Peg3* imprinted domain-controlling DMR (PCD) of cortical neurons

To explore the impact of neuronal tetraploidization on the methylation status of specific germline DMRs controlling the expression of neuron-expressed genes (26) we focused on the *Peg3* imprinted domain (Figure 1A) (8,9). For this analysis, we used cell nuclei from 2C and 4C neurons (identified by NeuN-specific immunoreactivity), isolated by FACS from the mouse cerebral cortex (Supplementary Figure S1), a structure that in adult mice contains around 3% of 4C neurons (11). Postnatal day 15 (P15) was chosen for this analysis, as *Peg3* is crucial for lactation and postnatal growth (27).

The proportion of methylated CpG sites within the PCD was analyzed in bisulfite-treated genomic DNA from either 2C or 4C neuronal nuclei. Bisulfite-treated genomic DNA from embryonic day 11 (E11) telencephalic neuroepithelium, a structure that contains mostly neural precursors, was used as a control. Bisulfite-modified genomic DNA was PCR amplified with primers specific for the *Peg3* domain (see Table 1), and the amplified bands were pyrosequenced using a specific internal primer (Figure 2A). This analysis

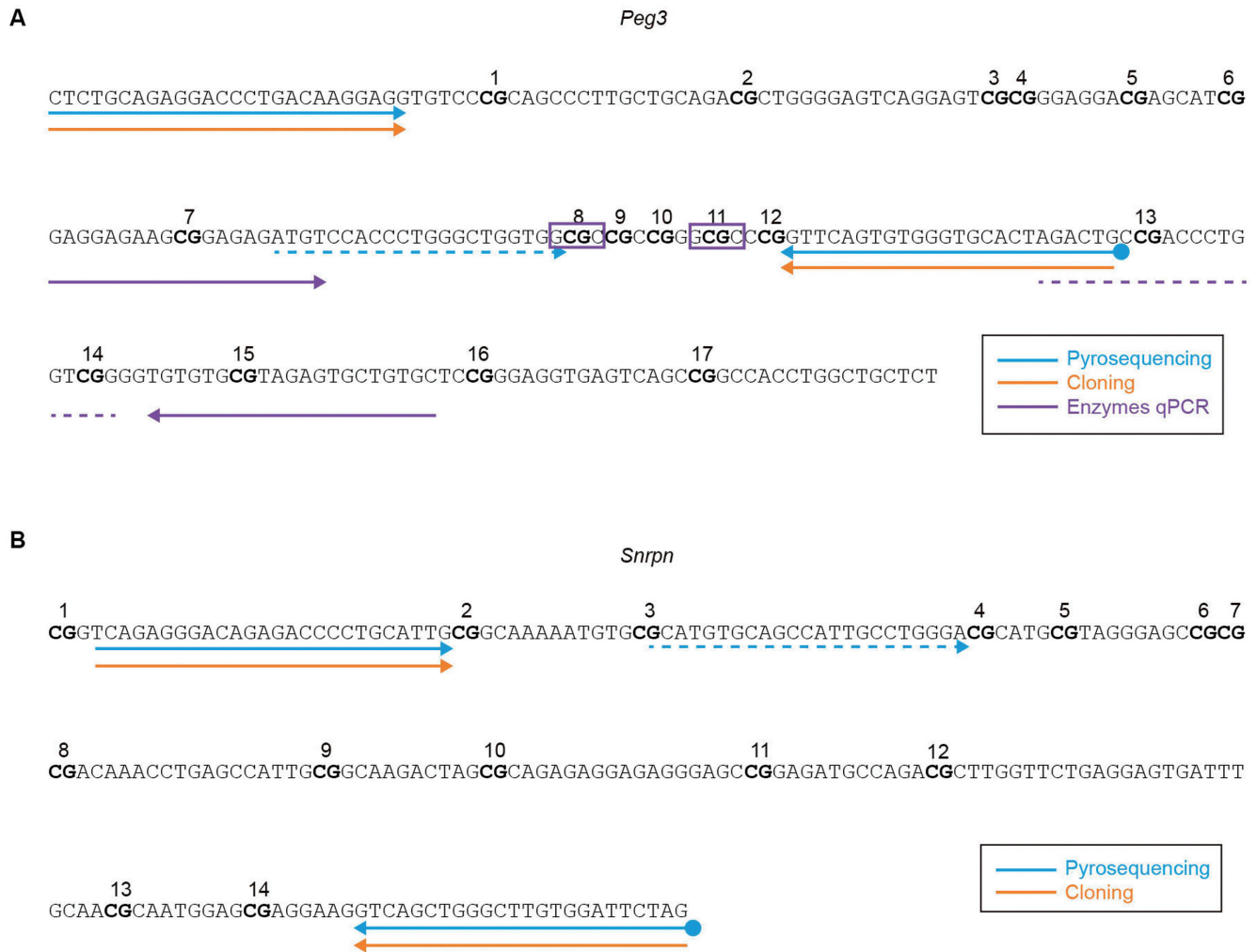


Figure 2. Genomic DNA sequences corresponding to the *Peg3*- and *Snrpn*-imprinting control regions analyzed in this study. (A and B) CpG sites are numbered and shown in bold, and HhaI restriction sites are indicated with purple rectangles. For qPCR amplification in (A), genomic DNAs were either cut with HhaI or left untreated. Solid purple arrows represent the sequences where the primers for qPCR were designed. Dashed purple line corresponds to the sequence used for the Taqman probe. For methylation analysis in both (A) and (B), genomic DNAs were bisulfite-converted and PCR amplified using modified primers (Cs substituted by Ts) corresponding to the sequences indicated by blue and orange solid arrows, and either cloned for Sanger sequencing [see Supplementary Figure S2] or subjected to pyrosequencing. In this latter case the downstream primers were 5'-biotinylated (solid circles), and pyrosequencing was performed with the primers indicated by dashed blue arrows (modified by substituting Cs by Ts in its sequence).

indicated that the global percentage of methylated CpGs in the PCD from 2C cortical neurons was significantly lower when compared with the E11 telencephalic neuroepithelium (Figure 3A). In addition, the methylation level observed in the PCD in the 4C cortical neurons was dramatically reduced with respect to 2C neurons (Figure 3A and B; see also Supplementary Figure S2A for Sanger sequencing). The decrease in CpG methylation observed in 4C neurons does not derive from their tetraploid status since the PCDs of 4C hepatocytes isolated from P20 livers show just a minimal reduction in the global levels of methylation compared with 2C (Figure 3C and D). The decrease in methylation observed in 4C neurons does not derive from partial genome duplication of the maternal chromosome in these neurons either. In support, whole genome amplification from 2C and 4C neuronal nuclei (22 000–27 000 on average) obtained from P15 cerebral cortex, followed by comparative genomic hy-

bridization (CGH), demonstrated that the genome of 4C cortical neurons is fully duplicated (Supplementary Figure S3). Targeted NGS of the *Peg3* genomic domain from 4C neuron-derived genomic DNA demonstrated that the sub-domain of the PCD analyzed in this study specifically contains a lower proportion of 5mCs (Supplementary Figure S4).

The reduction in the global percentage of methylated CpGs in 4C neurons is specific for the *Peg3* imprinted region as the *Snrpn* imprinting domain-controlling DMR (SCD) (8,28), also known to regulate a genomic domain (Figure 1B) comprising genes with neuronal expression (10), showed similar levels of global methylation in both 2C and 4C cortical neurons (Figure 3E). This conclusion was further supported by the analysis with specific primers (Figure 2) of individual CpGs within the SCD, either by pyrosequencing

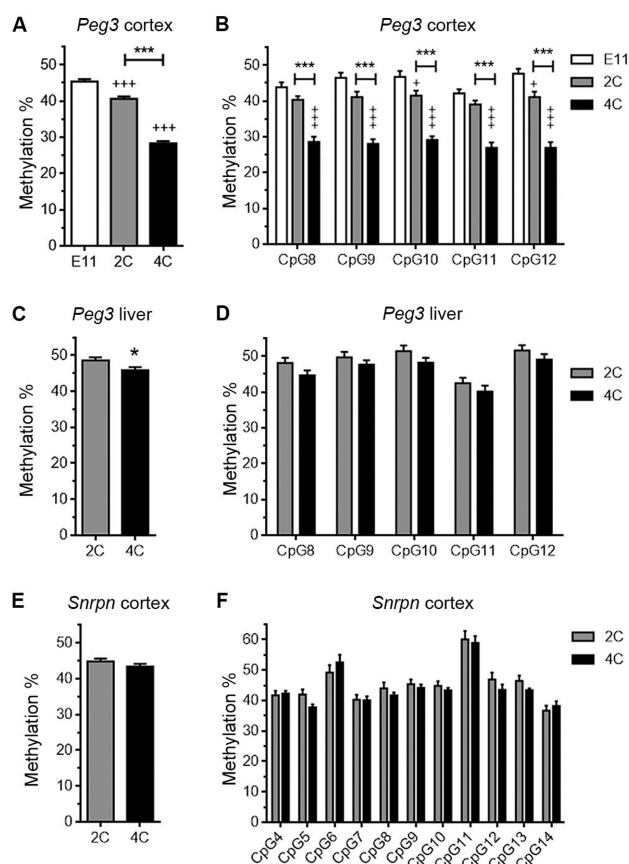


Figure 3. Reduced methylation levels in the PCD of cortical neurons. (A) Global pyrosequencing analysis of the CpG sites present in the PCD of E11 telencephalic neuroepithelial cells, and 2C and 4C cortical neurons. $+++P < 0.005$ (versus E11); $***P < 0.005$ (versus 2C). (B) Pyrosequencing analysis of individual CpG sites (numbered as in Figure 2A) in the PCD of E11 telencephalic neuroepithelial cells, and 2C and 4C cortical neurons. $+++P < 0.005$; $+P < 0.05$ (versus E11); $***P < 0.005$ (versus 2C). (C and E) Global pyrosequencing analysis of the CpG sites present in the PCD of 2C and 4C hepatocytes (C) and in the SCD of 2C and 4C cortical neurons (E). $*P < 0.05$ (versus 2C). (D and F) Pyrosequencing analysis of individual CpG sites (numbered as in Figure 2A) in the PCD of 2C and 4C hepatocytes (D) and in the SCD (numbered as in Figure 2B) of 2C and 4C cortical neurons (F).

quencing (Figure 3F) or Sanger sequencing (Supplementary Figure S2B).

Strand-specific reduction of methylation levels in the PCD of cortical neuron subpopulations

To identify the reason behind the reduction of CpG methylation levels within the PCD from cortical neurons, we used a variation of the hairpin-bisulfite PCR method (24). This technique is able to assess bisulfite protection patterns on complementary strands of individual DNA molecules. To this end, we used a hairpin linker that is ligated to BbvI-cleaved genomic DNA obtained from 2C and 4C neuronal nuclei (Figure 4A and B). This approach results in the attachment of the complementary strands of individual DNA molecules, which are subsequently denatured and subjected to bisulfite conversion, followed by PCR amplification with *Peg3*-specific primers upstream of the cleav-

age sequence (Figure 4A). After Sanger sequencing of the PCR fragments, the presence of bisulfite protected or unprotected cytosines in both DNA strands is analyzed (Figure 5B and C). This analysis demonstrated that 15% of the *Peg3*-controlling DMRs display strand-specific demethylation in the 4C neurons (Figure 5A). In addition, a reduced proportion of strand-specific demethylated DMRs (7%) was detected in the PCD of 2C neurons, an observation consistent with the comparatively lower proportion of methylated DNA in these latter cells (Figure 3A). In both types of neurons, demethylation affected the CpGs located on the *Peg3* sense strand. We therefore concluded that a portion of the CpGs from the PCD in the methylated chromosome, comprising at least CpGs #8–13 (Figure 4A), is asymmetrically methylated in the P15 cortical neurons. This analysis also indicated that CpG sites #14–17 are methylated on both strands (Figure 5A–C), as expected from the differential chromosome methylation associated with the imprinting process. Overall, these results indicate that the observed decrease of methylation in cortical neurons (Figure 3A), roughly mimicking the proportion of hemimethylated DMRs (Figure 5A), is likely due to the asymmetric removal of 5mCs in a subdomain of the PCD from the methylated chromosome. Interestingly, this hemimethylated subdomain contains recognition motifs for Pax-5, p53, RAR- β :RXR- α , RXR- α , Egr3, VDR, FOXP3 and PXR-1:RXR- α transcription factors, as evidenced by the PROMO tool available at <http://algen.lsi.upc.es>.

The hemimethylation status described above was confirmed through methylation-sensitive, restriction enzyme-based quantitative PCR (21,29,30) (see Experimental Procedures). HhaI was chosen for this analysis, as this restriction enzyme does not cut fully- and hemi-methylated DNA (31) (Figure 5D), as well as methylformylated DNA (Figure 5D). Genomic DNAs isolated from either 2C or 4C neuronal nuclei were either mock digested or digested with HhaI. DNAs were then amplified through real time PCR using oligonucleotides flanking the HhaI cleavage sites in the PCD [CpGs #8 and #11 (Figure 2A)]. This assay demonstrated that ~50% of genomic DNA remains uncleaved in both 2C and 4C cortical neurons (Figure 5E), thus indicating that half of the chromosomes are resistant to HhaI cleavage in the PCD. Overall, these results demonstrate that the reduction of methylation observed in the PCD is not due to full demethylation of the studied CpGs, but rather to partial, strand-specific demethylation of specific CpGs.

Reduced methylation in the PCD of cortical neurons is largely independent on the presence of 5fC

5fC can be found as a stable DNA modification in mammalian genomes (6), but it cannot be detected through bisulfite sequencing (32). To explore whether the hemimethylated CpG sites from the subdomain of the PCD could contain 5fC, genomic DNA from 2C and 4C neuronal nuclei was treated with BH or EA. These compounds are reducing agents that specifically transform 5fC into 5hmC (22,32), a molecule that is protected from bisulfite treatment. The reducing capacity of BH and EA was confirmed in our experiments, which included a control DNA containing 5fC in specific CpGs (20) (see Supplementary Figure S5). Neither

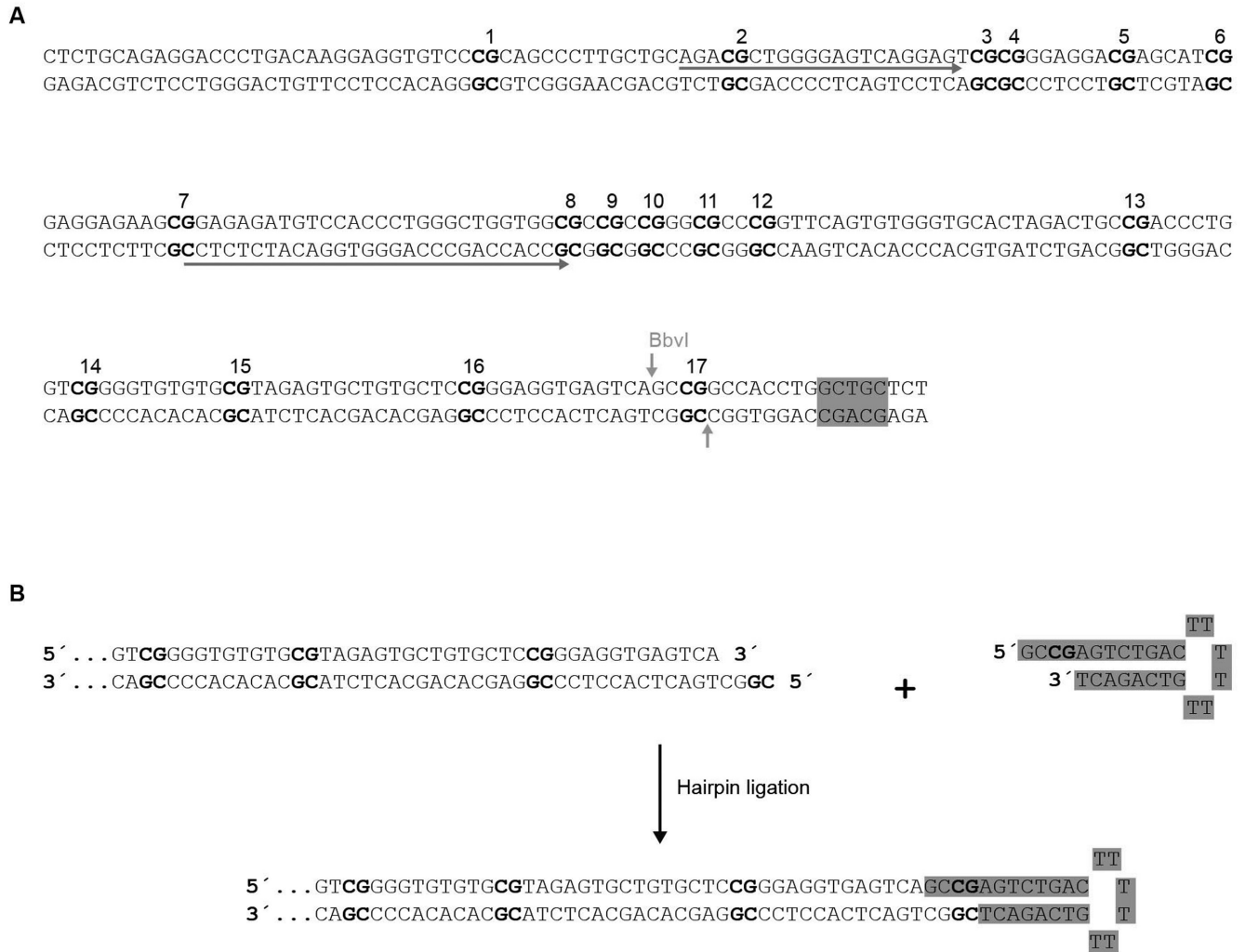


Figure 4. Scheme showing the strategy for hairpin-bisulfite PCR. (A) Genomic DNA sequence corresponding to the domain of the *Peg3* imprinting control region analyzed in this study. CpG sites are numbered and shown in bold. The BbvI recognition/cleavage site is indicated in gray. Sequences indicated by solid arrows correspond to the modified primers (Cs substituted by Ts) used for PCR. (B) Diagram illustrating the ligation of the BbvI-cleaved genomic DNA with the hairpin (labeled in gray). The resulting DNA is bisulfite-converted and subjected to PCR amplification with the oligonucleotides indicated in (A). The resulting amplification bands were cloned and Sanger sequenced (see Figure 5B and C).

BH nor EA treatment followed by bisulfite pyrosequencing was able to restore the proportion of global bisulfite-protected levels in the hypomethylated PCD subdomain from 2C neurons (Figure 6A,B). In contrast, BH was able to partially increase the global levels of bisulfite-protected CpGs in 4C neurons (Figure 6C), while EA treatment resulted in the tendency of a global increase of CpG methylation levels that was not statistically significant (Figure 6C). A BH- or EA-dependent increase in the proportion of bisulfite-protection in 4C neurons could not be confirmed when individual CpGs were analyzed either, although a statistically non-significant tendency could also be observed (Figure 6D). As a control, no significant increase was detected when similar analyses were performed with genomic DNA from E11 telencephalic neuroepithelium (Figure 6E and F), 2C hepatocytes (Figure 6G and H), and 4C hepatocytes (Figure 6I and J). Finally, no changes were observed when a similar analysis was performed in genomic

DNA from 2C neurons using primers specific for the SCD both globally (Figure 6K) and in individual CpG sites (Figure 6L). Overall, these results suggest that a subdomain of the PCD contains differentially methylformylated CpGs in a subpopulation of 4C cortical neurons. This pattern, which represents a novel epigenetic label, is likely to derive from chemical modification of the cytosines lying on one of the two DNA strands.

Maintenance of imprinting in the *Peg3* domain of cortical neurons

To study whether GpG hemimethylation participates in the maintenance of genomic imprinting in the *Peg3* domain, we focused on the *Peg3* gene as a read-out of the system. To this end, gene silencing in the maternal chromosome was analyzed using a single nucleotide polymorphism (SNP # rs31442449) located in intron 7 of the *Peg3* gene, characterized by a nucleotide change from T in C57BL/6J

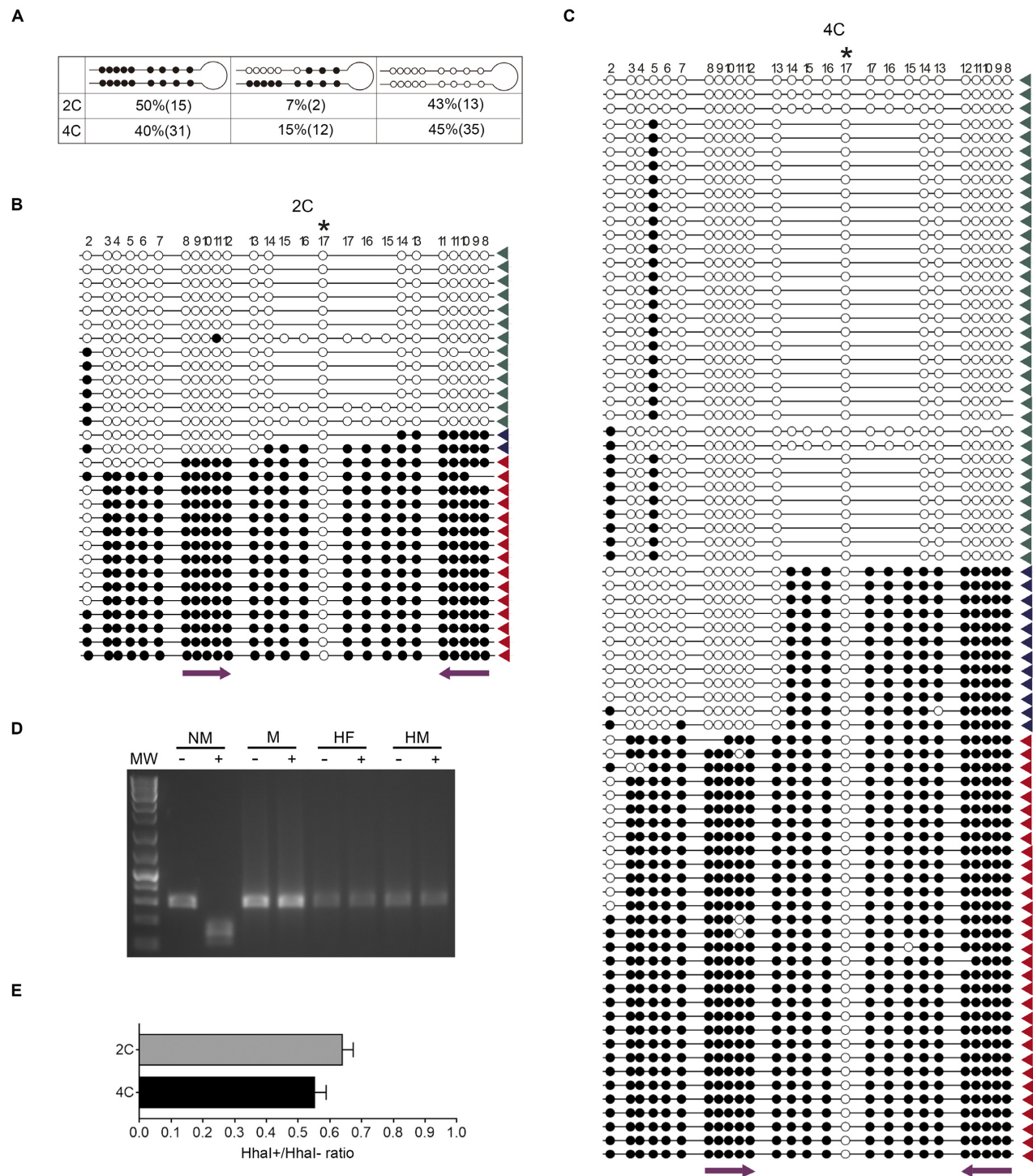


Figure 5. Reduced methylation levels in the PCD of cortical neurons is strand-specific. (A) Percentage and absolute number of hairpin-bisulfite PCR clones [see (B) and (C)] with the indicated bisulfite protection patterns. Unmethylated CpG sequences are indicated by open circles and methylated CpG sequences are indicated by solid circles. (B and C) Sanger sequences of the hairpin-bisulfite PCR clones from genomic DNA derived from 2C neurons (B) and 4C (C). Green arrowheads: clones with non-methylated CpG sites #8–12, blue arrowheads: clones with asymmetrically methylated CpG sites #8–12, red arrowheads: clones with bimethylated CpG sites #8–12. Non-methylated CpG sequences are indicated by open circles and methylated CpG sequences are indicated by solid circles. Asterisk indicates the CpG sequence from CpG site #17 that is included in the hairpin (see Figure 4A), and therefore is not methylated. CpG site #2 is shown as the oligonucleotide used for PCR amplification was degenerated. The localization of CpG sites #8–12 is indicated with purple arrows. (D) HhaI restriction assay using non-methylated DNA (NM), fully methylated DNA (M), DNA formylated in one strand and methylated in the other strand (HF), and DNA methylated in one strand (HM). -: Mock-treated DNA, +: HhaI-treated DNA, MW: DNA molecular weight marker. (E) Genomic DNA from 2C and 4C neuronal nuclei, undigested (HhaI-) or digested with HhaI (HhaI+), was subjected to real time PCR amplification using *Peg3*-specific primers flanking the restriction sites (see Figure 2A). Bars represent the HhaI+/HhaI- ratio average (\pm S.E.M.). Values were not significantly below 0.5 (one-tailed, Student's *t*-test).

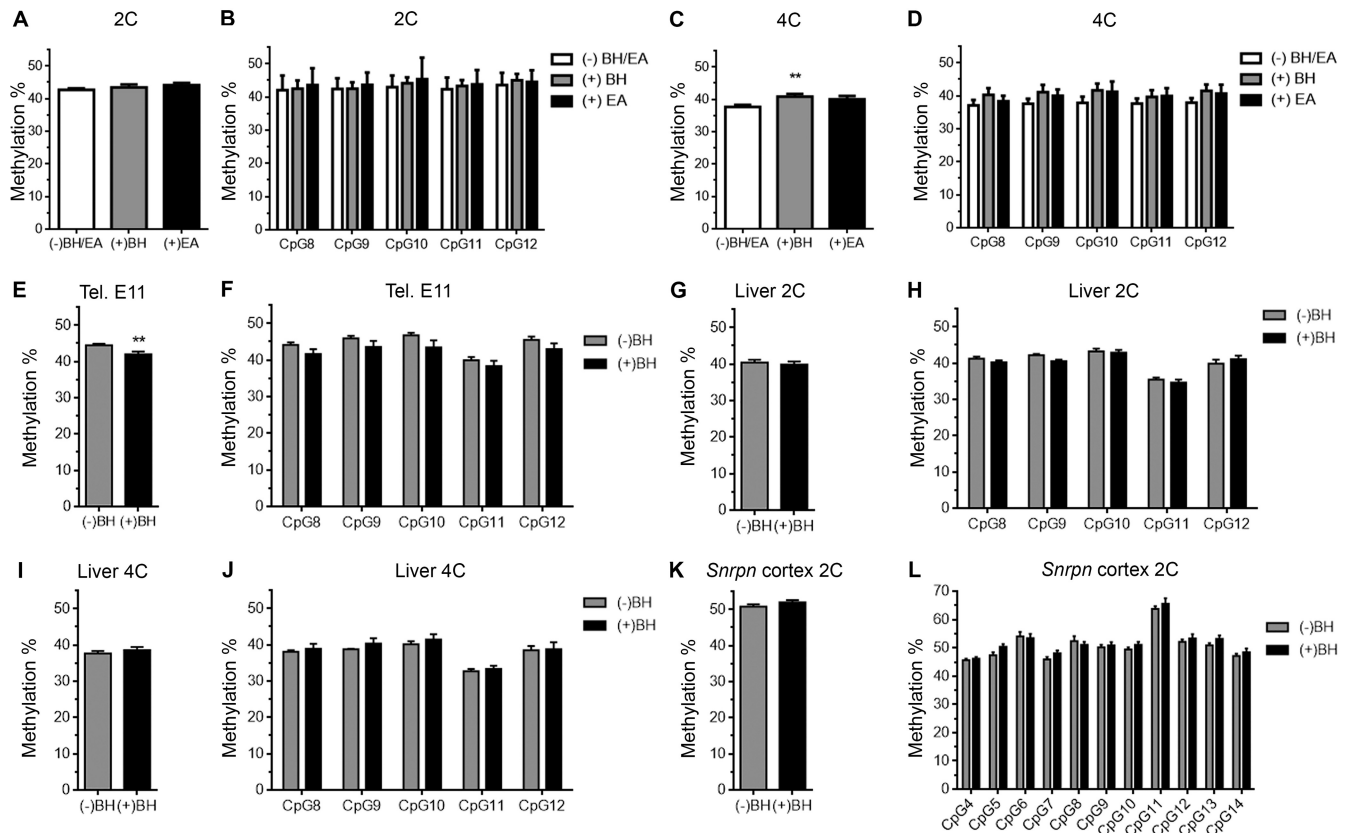


Figure 6. Reduced methylation levels in the PCD of 4C cortical neurons seems to be partially due to the presence of 5fC. (A) Global pyrosequencing analysis of the CpG sites present in the PCD, performed with untreated (white), BH- (gray), or EA-treated genomic DNA (black) from 2C cortical neurons. (B) Pyrosequencing analysis of individual CpG sites (numbered as in Figure 2A), performed with untreated (white), BH- (gray), or EA-treated genomic DNA (black) from 2C cortical neurons. (C) Global pyrosequencing analysis of the CpG sites present in the PCD, performed with untreated (white), BH- (gray), or EA-treated genomic DNA (black) from 4C cortical neurons. (D) Pyrosequencing analysis of individual CpG sites (numbered as in Figure 2A), performed with untreated (white), BH- (gray), or EA-treated genomic DNA (black) from 4C cortical neurons. (E, G and I) Global pyrosequencing analysis, performed with genomic DNA treated either with BH (black) or left untreated (gray), of the CpG sites present in the PCD of E11 telencephalic neuroepithelium (E), 2C hepatocytes (G), and 4C hepatocytes (I). (F, H and J) Pyrosequencing analysis, performed with genomic DNA treated either with BH (black) or left untreated (gray), of individual CpG sites (numbered as in Figure 2A) present in the PCD of E11 telencephalic neuroepithelium (F), 2C hepatocytes (H) and 4C hepatocytes (J). (K) Global pyrosequencing analysis, performed with genomic DNA treated either with BH (black) or left untreated (gray), of the CpG sites present in the SCD of 2C cortical neurons. (L) Pyrosequencing analysis, performed at P15 with genomic DNA treated either with BH (black) or left untreated (gray), of individual CpG sites (numbered as in Figure 2B) present in the SCD of 2C cortical neurons. ** $P < 0.01$.

mice to C in DBA/2J mice (Figure 7A). Female C57BL/6J mice and male DBA/2J mice were mated to generate hybrid mice heterozygous for the T/C polymorphism within the maternal/paternal chromosomes, respectively (Figure 7B). Nascent mRNA (enriched in unspliced transcripts) was then obtained from 2C and 4C neuronal nuclei, isolated by FACS from the cerebral cortex of the hybrid mice at P15. In addition, nascent RNA was also obtained from telencephalic neuroepithelial cell nuclei of E11 heterozygous embryos. These mRNAs were not contaminated by genomic DNA as no specific amplification was detected after 100 PCR cycles with the primers used in the subsequent analysis (Figure 7C). Furthermore, the results indicated absence of maternal expression (Figure 7 and Supplementary Figure S6), therefore ruling out the existence of genomic contamination. Pyrosequencing analysis of the cDNAs derived from these mRNAs indicated that the *Peg3* gene remains silenced in the maternal chromosome, since only background signal was obtained for the T-containing *Peg3* sequence in

both 2C (Figure 7E) and 4C (Figure 7F) cortical neurons, which was similar to that observed in the E11 neuroepithelium (Figure 7D). The silencing of the maternal *Peg3* gene in the E11 neuroepithelium and 2C neurons was confirmed by additional pyrosequencing using another primer that was closer to the SNP (Figure 7G,H). These results were confirmed in reciprocal crosses (female DBA/2J mice mated to male C57BL/6J mice) (Supplementary Figure S6).

DISCUSSION

In this study, we have shown that the methylation pattern of the DMR controlling the murine *Peg3* imprinted domain can be altered in cortical neurons. In subpopulations of these cells, the normal pattern, characterized by symmetrical methylation of the constituting CpG sites that are present in the maternal allele, is changed to a strand-specific, hemimethylated pattern observed in a small segment of the *Peg3*-controlling DMR. This conclusion was reached after using five different approaches: bisulfite

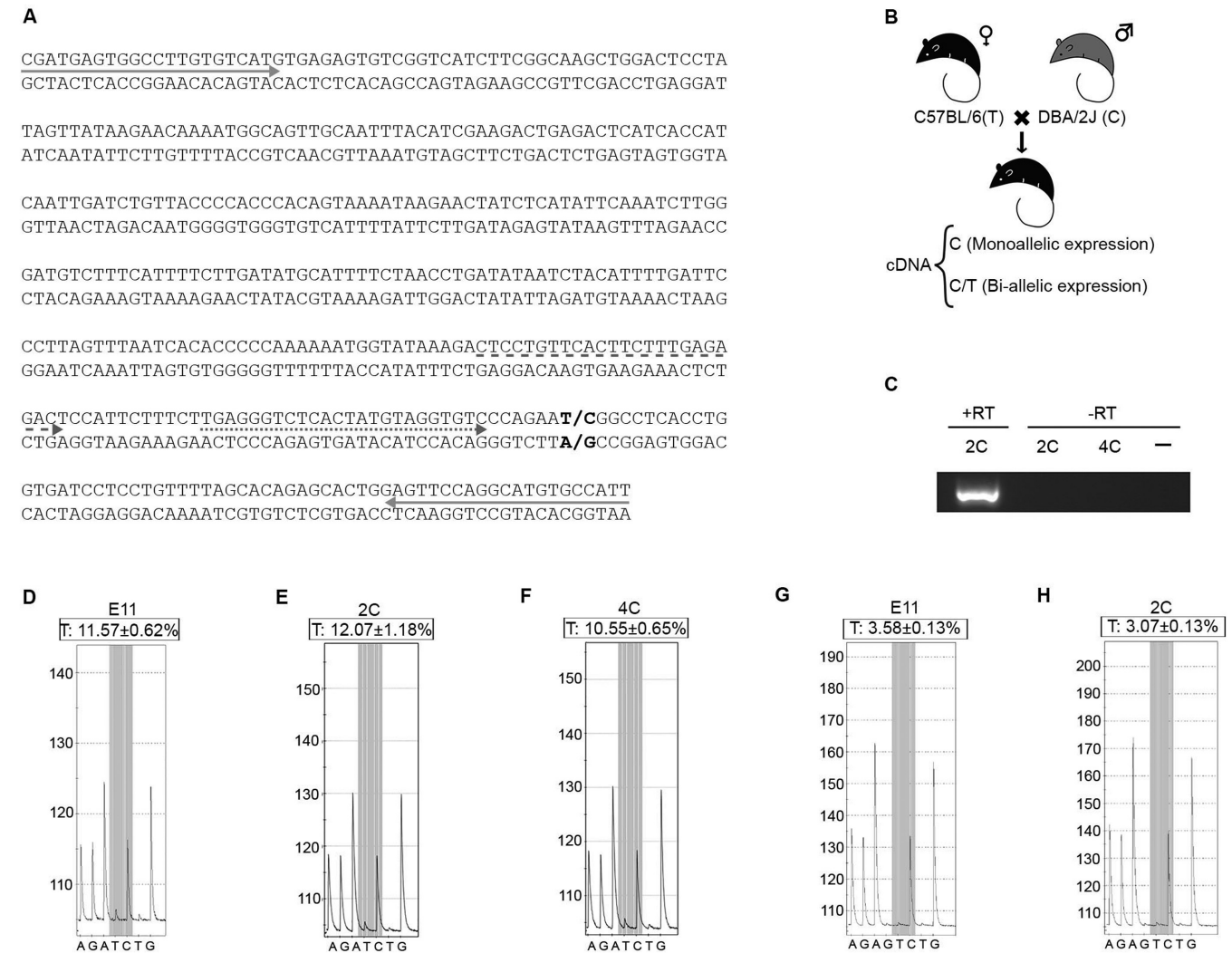


Figure 7. PCD hemimethylation can maintain the monoallelic expression pattern of the *Peg3* imprinted gene. (A) Partial cDNA sequence from the *Peg3* unspliced transcript (amplified using the primers indicated with solid arrows) is shown. Specific primers [dashed arrow: primer used for (D–F); dotted arrow: primer used for (G and H)] were used to pyrosequence the region containing SNP # rs31442449 (shown in bold). (B) Scheme of the breeding design using mice with a SNP in the *Peg3* gene, and the expected cDNAs in case of either mono or biallelic expression. (C) PCR amplification with the primers indicated with solid arrows in (A) using mock reverse transcribed (–RT) or reverse transcribed (+RT) mRNAs from 2C or 4C neuronal nuclei or water (–). 50 PCR cycles were given to the +RT cDNA and 50 + 50 cycles were given to the –RT reaction. (D–F) Representative pyrograms from the cDNA sequence in the E11 telencephalic neuroepithelium (D), 2C (E), and 4C (F) cortical neurons, using the primer indicated with a dashed arrow in (A). The SNP (T/C) is shown in grey. T from position 6 is a negative control. Two different cDNAs were used for each analysis. Mean of the proportion of T4 (±S.E.M.) is shown. (G and H) Representative pyrograms from the cDNA sequence in the E11 telencephalic neuroepithelium (G) and 2C (H) cortical neurons, using the primer indicated with a dotted arrow in (A). The SNP (T/C) is shown in grey. G from position 4 (G4) and T7 are negative controls. Mean of the proportion of T7 (±S.E.M.) is shown.

Sanger sequencing, bisulfite pyrosequencing, methylation-sensitive DNA cleavage followed by real time PCR, hairpin bisulfite PCR, and targeted NGS analysis. Since it takes place in a small region of the PCD inserted within a fully methylated domain (Figure 5A–C), we believe this strand-specific modification affects the maternal allele.

Genomic imprinting prevents parthenogenesis (33) and polyploidy (34,35) in mammalian embryos by regulating genes that are usually required for normal development (36). In addition, diploid/triploid mosaicism in somatic tissues may cause genetic syndromes with aberrant genomic imprinting (37). In this study, we have shown that the imprinting process in 4C cortical neurons is not altered. This

indicates that 4C neurons generated during normal development (11,12) are unlikely to have mechanisms that disrupt imprinting.

Strand-specific CpG hemimethylation is not specific for tetraploid neurons since it can also be detected in diploid neurons. In addition, not all tetraploid neurons show this modification. The most likely explanation is that this epigenetic mark affects to a subset of cortical neurons, which are mainly tetraploid.

Our study challenges the current view that the methylation pattern of germline DMRs cannot be altered in somatic cells as they can contain hemimethylated and probably methylformylated subdomains. Further studies are needed

to evaluate the biological impact of the modifications observed in the PCD of specific cortical neurons. Our results indicate that the CpG hemimethylation within the murine PCD specifically affects one of the two DNA strands. This suggests the existence of a novel molecular mechanism regulating the transition from fully methylated CpGs to strand-specific formylated CpGs. We note that the *Peg3*-controlling DMR is located within the *Peg3* gene promoter, and that promoters are known to be enriched in 5fC (38). As also happens with full CpG formylation (39), CpG methylformylation might alter the DNA structure of the maternal *Peg3* promoter. This may result in changes of *Peg3* expression from the paternal chromosome due to the complex regulation that seems to affect to the PCD (40), while maintaining the imprinting capacity of the *Peg3* domain. In addition, we cannot rule out that specific 5fC interactors with defined functions (41) or putative proteins interacting with hemimethylated CpGs can participate in the regulation of the PCD. The residual presence of 5fC within the murine PCD of 4C cortical neurons could also be interpreted as an intermediate state of the strand-specific demethylation process. Uncovering the molecular mechanism leading to strand-specific demethylation will facilitate our understanding of the function of this novel epigenetic label.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Sergio Casas-Tintó for discussion about this work, Mark Strunk for his technical help, Simon Heath and Angelika Merkel for their help with the bioinformatic analysis, and Chaska Walton for scientific discussion and the final English version of the Manuscript.

FUNDING

Fundación Ramón Areces [CIVP16A1815]; Spanish Ministry of Economy and Competitiveness [SAF2015-68488-R]; PhD contract from the Spanish Ministry of Economy and Competitiveness (to I.P.P.). Funding for open access charge: Spanish Ministry of Economy and Competitiveness [SAF2015-68488-R].

Conflict of interest statement. None declared.

REFERENCES

- Wan, L. and Bartolomei, M.S. (2008) Regulation of imprinting in clusters: noncoding RNAs versus insulators. *Adv. Genet.*, **61**, 207–223.
- Nowak, K., Stein, G., Powell, E., He, L.M., Naik, S., Morris, J., Marlow, S. and Davis, T.L. (2011) Establishment of paternal allele-specific DNA methylation at the imprinted mouse *Gtl2* locus. *Epigenetics*, **6**, 1012–1020.
- Leonhardt, H., Page, A.W., Weier, H.U. and Bestor, T.H. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell*, **71**, 865–873.
- Constância, M., Pickard, B., Kelsey, G. and Reik, W. (1998) Imprinting mechanisms. *Genome Res.*, **8**, 881–900.
- Suzuki, M.M. and Bird, A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.*, **9**, 465–476.
- Bachman, M., Uribe-Lewis, S., Yang, X., Burgess, H.E., Iurlaro, M., Reik, W., Murrell, A. and Balasubramanian, S. (2015) 5-Formylcytosine can be a stable DNA modification in mammals. *Nat. Chem. Biol.*, **11**, 555–557.
- Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C. and Zhang, Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
- Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H. and Trasler, J.M. (2002) Methylation dynamics of imprinted genes in mouse germ cells. *Genomics*, **79**, 530–538.
- Huang, J.M. and Kim, J. (2009) DNA methylation analysis of the mammalian PEG3 imprinted domain. *Gene*, **442**, 18–25.
- Schmauss, C., Brines, M.L. and Lerner, M.R. (1992) The gene encoding the small nuclear ribonucleoprotein-associated protein N is expressed at high levels in neurons. *J. Biol. Chem.*, **267**, 8521–8529.
- López-Sánchez, N. and Frade, J.M. (2013) Genetic evidence for p75^{NTR}-dependent tetraploidy in cortical projection neurons from adult mice. *J. Neurosci.*, **33**, 7488–7500.
- Morillo, S.M., Escoll, P., de la Hera, A. and Frade, J.M. (2010) Somatic tetraploidy in specific chick retinal ganglion cells induced by nerve growth factor. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 109–114.
- Shirazi Fard, S., Jarrin, M., Boije, H., Fillon, V., All-Eriksson, C. and Hallböök, F. (2013) Heterogenic final cell cycle by chicken retinal Lim1 horizontal progenitor cells leads to heteroploid cells with a remaining replicated genome. *PLoS One*, **8**, e59133.
- Morillo, S.M., Abanto, E.P., Román, M.J. and Frade, J.M. (2012) Nerve growth factor-induced cell cycle reentry in newborn neurons is triggered by p38^{MAPK}-dependent E2F4 phosphorylation. *Mol. Cell. Biol.*, **32**, 2722–2737.
- Ovejero-Benito, M.C. and Frade, J.M. (2013) Brain-derived neurotrophic factor-dependent cdk1 inhibition prevents G2/M progression in differentiating tetraploid neurons. *PLoS One*, **8**, e64890.
- Ovejero-Benito, M.C. and Frade, J.M. (2015) p27^{Kip1} participates in the regulation of endoreplication in differentiating chick retinal ganglion cells. *Cell Cycle*, **14**, 2311–2322.
- Kaczmarek, L. and Chaudhuri, A. (1997) Sensory regulation of immediate-early gene expression in mammalian visual cortex: implications for functional mapping and neural plasticity. *Brain Res. Brain Res. Rev.*, **23**, 237–256.
- Knapka, E. and Kaczmarek, L. (2004) A gene for neuronal plasticity in the mammalian brain: *Zif268/Egr-1/NGFI-A/Krox-24/TIS8/ZENK*? *Prog. Neurobiol.*, **74**, 183–211.
- Kaufman, M.H. (1992) *The Atlas of Mouse Development*. Academic Press Inc., San Diego.
- Booth, M.J., Marsico, G., Bachman, M., Beraldi, D. and Balasubramanian, S. (2014) Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nat. Chem.*, **6**, 435–440.
- Hashimoto, K., Kokubun, S., Itoi, E. and Roach, H.I. (2007) Improved quantification of DNA methylation using methylation-sensitive restriction enzymes and real-time PCR. *Epigenetics*, **2**, 86–91.
- Song, C.X., Szulwach, K.E., Dai, Q., Fu, Y., Mao, S.Q., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H. et al. (2013) Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, **153**, 678–691.
- Schuyler, R.P., Merkel, A., Raineri, E., Altucci, L., Vellenga, E., Martens, J.H., Pourfarzad, F., Kuijpers, T.W., Burden, F., Farrow, S. et al. (2016) Distinct trends of DNA methylation patterning in the innate and adaptive immune systems. *Cell Rep.*, **17**, 2101–2111.
- Laird, C.D., Pleasant, N.D., Clark, A.D., Sneed, J.L., Hassan, K.M., Manley, N.C., Vary, J.C. Jr, Morgan, T., Hansen, R.S. and Stöger, R. (2004) Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 204–209.
- Kumaki, Y., Oda, M. and Okano, M. (2008) QUMA: quantification tool for methylation analysis. *Nucleic Acids Res.*, **36**, W170–W175.
- Relaix, F., Weng, X., Marazzi, G., Yang, E., Copeland, N., Jenkins, N., Spence, S.E. and Sassoon, D. (1996) Pw1, a novel zinc finger gene implicated in the myogenic and neuronal lineages. *Dev. Biol.*, **177**, 383–396.

27. Curley,J.P., Barton,S., Surani,A. and Keverne,E.B. (2004) Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. *Proc. Biol. Sci.*, **271**, 1303–1309.
28. Miyazaki,K., Mapendano,C.K., Fuchigami,T., Kondo,S., Ohta,T., Kinoshita,A., Tsukamoto,K., Yoshiura,K., Niikawa,N. and Kishino,T. (2009) Developmentally dynamic changes of DNA methylation in the mouse Snurf/Snrpn gene. *Gene*, **432**, 97–101.
29. Hua,D., Hu,Y., Wu,Y.Y., Cheng,Z.H., Yu,J., Du,X. and Huang,Z.H. (2011) Quantitative methylation analysis of multiple genes using methylation-sensitive restriction enzyme-based quantitative PCR for the detection of hepatocellular carcinoma. *Exp. Mol. Pathol.*, **91**, 455–460.
30. Bai,L., Yan,P., Cao,X., Jia,L., Zhang,C. and Guo,D. (2015) Methylation-sensitive restriction enzyme nested real time PCR, a potential approach for sperm DNA identification. *J. Forensic Leg. Med.*, **34**, 34–39.
31. Bird,A.P. (1978) Use of restriction enzymes to study eukaryotic DNA methylation: II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J. Mol. Biol.*, **118**, 49–60.
32. Plongthongkum,N., Diep,D.H. and Zhang,K. (2014) Advances in the profiling of DNA modifications: cytosine methylation and beyond. *Nat. Rev. Genet.*, **15**, 647–661.
33. Spindle,A., Sturm,K.S., Flannery,M., Meneses,J.J., Wu,K. and Pedersen,R.A. (1996) Defective chorioallantoic fusion in mid-gestation lethality of parthenogenone ⇌ tetraploid chimeras. *Dev. Biol.*, **173**, 447–458.
34. Devriendt,K. (2005) Hydatidiform mole and triploidy: the role of genomic imprinting in placental development. *Hum. Reprod. Update*, **11**, 137–142.
35. Bacquet,C., Imamura,T., Gonzalez,C.A., Conejeros,I., Kausel,G., Neildes-Nguyen,T.M., Paldi,A. and Gallardo,M.H. (2008) Epigenetic processes in a tetraploid mammal. *Mamm. Genome*, **19**, 439–447.
36. Biniszkiewicz,D., Gribnau,J., Ramsahoye,B., Gaudet,F., Eggan,K., Humpherys,D., Mastrangelo,M.A., Jun,Z., Walter,J. and Jaenisch,R. (2002) Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Mol. Cell. Biol.*, **22**, 2124–2135.
37. Rittinger,O., Kronberger,G., Pfeifenberger,A., Kotzot,D. and Fauth,C. (2008) The changing phenotype in diploid/triploid mosaicism may mimic genetic syndromes with aberrant genomic imprinting: follow up in a 14-year-old girl. *Eur. J. Med. Genet.*, **51**, 573–579.
38. Neri,F., Incarnato,D., Krepelova,A., Rapelli,S., Anselmi,F., Parlato,C., Medana,C., Dal Bello,F. and Oliviero,S. (2015) Single-base resolution analysis of 5-formyl and 5-carboxyl cytosine reveals promoter DNA methylation dynamics. *Cell Rep.*, **10**, 674–683.
39. Ngo,T.T., Yoo,J., Dai,Q., Zhang,Q., He,C., Aksimentiev,A. and Ha,T. (2016) Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability. *Nat. Commun.*, **7**, 10813.
40. Perera,B.P. and Kim,J. (2016) Sex and tissue specificity of Peg3 promoters. *PLoS One*, **11**, e0164158.
41. Song,J. and Pfeifer,G.P. (2016) Are there specific readers of oxidized 5-methylcytosine bases? *Bioessays*, **38**, 1038–1047.

Chapter 4

Flow Cytometric Quantification, Isolation, and Subsequent Epigenetic Analysis of Tetraploid Neurons

Noelia López-Sánchez, Iris Patiño-Parrado, and José María Frade

Abstract

Different forms of genomic mosaicism can be detected in vertebrate neurons, including copy number variation, L1 transposition, and aneuploidy. In addition, some populations of vertebrate neurons can also show double the normal amount of DNA, a condition referred to as somatic tetraploidy. These neurons are generated during early stages of development, as they migrate to their adult locations in the adult nervous system, and constitute subpopulations of projection neurons. We have previously shown that neuronal tetraploidy can be characterized by flow cytometry using isolated cell nuclei from different mammalian and avian structures. In this chapter, we describe this procedure using a different model system: the rhombencephalic derivatives from adult zebrafish. In addition, tetraploid neuronal nuclei can be isolated by fluorescence-activated cell sorting and their genomic DNA used for further analyses, either used directly or after whole-genome amplification. Here we show as an example how to perform epigenetic analyses to characterize CpG methylation in differentially methylated regions controlling the *Rasgrf1*-imprinting domain in mice.

Key words Cell sorting, FACS, Tetraploid neuron, Cell nuclei, Zebrafish, Genomic imprinting, DNA methylation, Pyrosequencing

1 Introduction

Polyploidy can be defined as the presence in a cell of multiple copies of the haploid chromosome complement, which is usually referred to as N . This condition can affect all the cells in the body (i.e., “germinal polyploidy”), a strategy adopted by several organisms to increase their size. This type of polyploidy is widespread in plants [1–3], and it can often be observed in animals [4], including protozoa [5], insects [6], fishes [7], amphibians [8, 9], and reptiles [10].

With the occurrence of multicellularity and the acquisition of the soma, specific cell types or tissues can show changes in the amount of DNA proportional to the haploid complement, thus

Noelia López-Sánchez and Iris Patiño-Parrado are contributed equally to this work.

José María Frade and Fred H. Gage (eds.), *Genomic Mosaicism in Neurons and Other Cell Types*, Neuromethods, vol. 131, DOI 10.1007/978-1-4939-7280-7_4, © Springer Science+Business Media LLC 2017

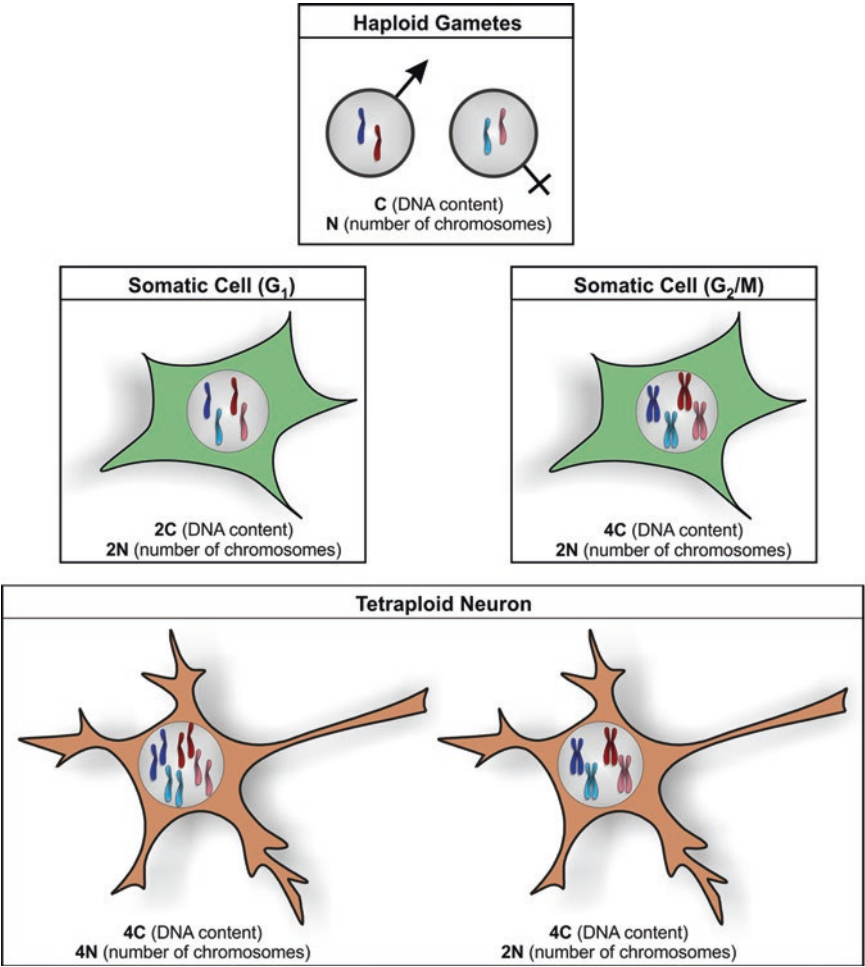


Fig. 1 *C* value and *N* value in germinal cells (gametes), somatic cells in G₁ and G₂/M phases of the cell cycle, and in tetraploid neurons. *C* value refers to the amount of DNA present in the gametes whereas *N* value denotes the chromosomal complement that characterizes the gametes. Since chromosomes may contain one or two chromatids the *C* values is duplicated during G₂ while the *N* value is unaltered as the cell progresses through the cell cycle. Since neurons cannot divide, the number of chromosomes of those with 4C DNA content cannot be identified and they are referred to as somatic tetraploid neurons *sensu lato*

resulting in “somatic polyploidy.” This term is used in a generic meaning since, in the absence of mitosis, the identification of the actual number of chromosomal sets is not feasible. In this case, somatic polyploid cells are usually referred to the amount of DNA that is present in their nuclei (measured in terms of the *C* value, i.e., the amount of DNA present in the haploid complement). It is important not to confound the *C* value with the *N* value: while the former increases during S phase and remains duplicated during G₂, *N* is unaltered as the cell progresses through the cell cycle (Fig. 1).

Somatic polyploidy can affect different tissues in both vertebrates and invertebrates, where it may serve as a means to increase

cell size and to promote gene expression and metabolic activity [11], which could facilitate better adaptation to metabolic and/or genotoxic stress and a lower sensibility to apoptotic stimuli [12–16]. In invertebrates, somatic polyploidy can be detected for instance in the glial cells making the blood–brain barrier [17] and in oocyte nurse cells [18] in *Drosophila*, and in Malpighian tubules and silk glands in the flour moth [19]. In mammals, a number of cell types have been shown to be polyploid, including megakaryocytes [20], trophoblast giant cells [21], hepatocytes [22], keratinocytes [23], vascular smooth cells [24], and cardiomyocytes [25].

As other tissues, the nervous system can also contain somatic polyploid cells. In invertebrates, several examples of somatic polyploid neurons have been described [26, 27]. The augment in the size of these neurons is likely required for increasing the speed at which action potentials move along their axons [28]. In some instances, the levels of ploidy may be really enormous as it is the case for nuclei of giant neurons from *Aplysia californica*, which has been shown to contain 200,000-fold the normal amount of haploid DNA (i.e., 200,000C) [29].

1.1 Somatic Tetraploidy in Vertebrate Neurons

For decades, it has been assumed that the vertebrate nervous system is homogenous from a genomic point of view, being constituted by diploid neurons (i.e., neurons showing 2C DNA content) [30]. This belief was challenged in the 1968 when Lapman [31] proposed that Purkinje cells contain tetraploid nuclei. This question has been readdressed, and current evidence indicates that a small proportion of neurons in the vertebrate nervous system contain tetraploid neurons (i.e., neurons with 4C DNA content) (Fig. 1). By using state-of-the-art techniques, Thomas Arendt's laboratory has shown that around 1% of neurons are tetraploid in the cerebral cortex of normal individuals [32]. In addition, we and others have demonstrated that tetraploid neurons exist in the normal vertebrate retina [33–35], as well as in the chick forebrain and mouse cerebral cortex [36]. In the chick retina, tetraploid neurons constitute subpopulations of horizontal cells [34, 35] and RGCs, the latter innervating lamina F at the *stratum-griseum-et-fibrosum-superficiale* of the tectal cortex [33]. These neurons are generated during the early stages of retinal development, soon after they acquire initial neuronal markers. In this regard, a subset of migrating RGCs expressing the transcription factor E2F1, and lacking Rb protein, undergo S phase and remain in a permanent G2-like state, due to the inactivation of Cdk1 in these cells [37]. In these cells, p27^{Kip1} expression prevents further round of replication, thus keeping these neurons in a tetraploid state [38]. Therefore, endoreduplication, but not alternative mechanisms such as aneuploidy or cell fusion [39, 40], seems to represent the major mechanism generating tetraploid RGCs in the vertebrate nervous system.

1.2 Methods to Characterize Tetraploidy in Vertebrate Neurons

As indicated above, the initial attempts for DNA content quantification in vertebrate neurons took place in the late 1960s [31]. At this early period, DNA quantification was performed by using a histochemical method, the Feulgen stain [41]. This colorimetric procedure is based on the acid hydrolysis of DNA, which produces free aldehyde groups that are subsequently detected [42]. Then, the intensity of the Feulgen reaction, which is proportional to the DNA concentration, can be measured using microdensitometric procedures [43]. Although many laboratories were focused on the analysis of neuronal tetraploidy using this technology [44–46], no clear conclusion was reached at that time due to the limitations inherent to the Feulgen technique applied to cells with low levels of polyploidy [47].

More recently, a number of methods based on fluorescent labeling of DNA have been developed to identify and characterize neuronal tetraploidy. These methods allow the identification through fluorescent labeling of the cell types that become tetraploid in the vertebrate nervous system. In Chap. 5, Thomas Arendt and colleagues describe a procedure for DNA quantification in single cells in brain slices based on image cytometry. This method relies on laser scanning cytometry-based technology applied to tissue sections labeled with fluorescent DNA dyes such as 4',6-diamidino-2-phenylindole (DAPI) or propidium iodide (PI) [48], which stoichiometrically bind to DNA. By using this method, these authors showed a positive correlation between the fluorescence intensity signal and both the number of dots detected with chromogenic in situ hybridization (CISH) using chromosome-specific probes and the amplification levels of *alu* repeats obtained by real-time PCR amplification in microdissected neurons [32].

A robust method for the analysis of neuronal tetraploidy is flow cytometry, a technique that allows the analysis of DNA content, protein expression, and other functional parameters in isolated cells or organelles in suspension [49–52]. An initial attempt for DNA content analysis in cerebral tissue using this technique was performed using fluorescent labeling of DNA through the acriflavine-Feulgen method [53]. This protocol uses formalin-fixed cell nuclei subjected to the acidic treatment inherent to the Feulgen stain, which includes acriflavine to provide fluorescence. This primitive protocol leads to a high frequency of nuclear doublets that cannot be removed from the analysis, thus impeding reliable quantification of 4C nuclei. Furthermore, the acriflavine-Feulgen method is too aggressive to allow parallel labeling to identify the nature of the tetraploid cell nuclei.

An important improvement for DNA analysis using flow cytometry derived from the use of fluorescent DNA dyes, such as PI [54]. The subsequent development of methods to remove doublets from the analysis, such as the standard pulse processing method [55] which we routinely use in our studies (see [36], for

example), was an additional advance for the quantitative power of flow cytometry. Many studies have been performed since the development of this improved methodology. Embryonic neurons can be isolated through controlled trypsin digestion of developing neural tissues following a protocol similar to that used for cell culture. Cell suspensions are then subjected to a number of steps including ethanol fixation, cell immunolabeling with specific neuronal markers, RNA removal with RNase, and stoichiometric DNA labeling with PI, prior to flow cytometric analysis [56]. In the past, our laboratory has used this procedure for the analysis of neuronal tetraploidy in cell suspensions from different neural tissues obtained from the chick embryo [33, 57].

A further improvement to perform quantitative analyses by flow cytometry was the use of fresh cell nuclei isolated by tissue grinding [58]. This approach eliminates the need for fluorescent DNA dyes to cross a permeabilized plasma membrane and reduces nonspecific dye binding to cytoplasmic components, allowing very accurate DNA cell cycle analysis. Most protocols for cell nuclei analysis by flow cytometry use fixed cell nuclei, either by ethanol or formaldehyde [59, 60], although examples of flow cytometric analysis using fresh cell nuclei are also available in the literature [61]. It is important to point out that the analysis of DNA content with fresh cell nuclei avoids the disadvantages of formaldehyde, which triggers chromatin cross-linking that impairs stoichiometric analysis of DNA content [62], and of ethanol, which may lead to cell or organelle aggregation [63].

In our laboratory we have developed a flow cytometric method for DNA content quantification that uses fresh cell nuclei from neural tissues obtained with minimal manipulation [36, 64, 65]. This protocol, which is described in detail later in this chapter, results in narrow coefficients of variation that allows accurate DNA quantification in our samples. For additional information, see Chap. 2 in which this technology is used for nuclear DNA content analysis, nucleotide sequencing, and interphase FISH.

1.3 Methods for the Isolation of the Genome of Tetraploid Neurons

A simple way to isolate cell nuclei from hyperploid neurons relies on the use of laser capture microdissection from tissue sections, as shown by Mosch et al. [32]. Nevertheless, this method is time consuming, thus precluding its use for the analysis of a high number of cell nuclei. Therefore, other preparative procedures based on flow cytometry have been developed. Fluorescence-activated cell sorting (FACS) is a technical development of flow cytometry that enables sorting of a mixture of organelles or cells into two or more fractions using the scatter and fluorescence signals of each particle [66]. This methodology has been used by our laboratory to isolate neural precursors in different stages of the cell cycle [67] or live neurons out of cortical cells [68].

Immunolabeling of fresh cell nuclei with antibodies against nuclear markers specific for neuronal subtypes allows the isolation of rather pure neuronal cell populations through FACS. This approach yields sufficient amount of genomic DNA from cell populations enriched in tetraploid neurons, which can be used for genomic analyses either directly or after whole-genome amplification. Nascent RNA can also be obtained from these nuclear preparations, thus resulting in a convenient method for the analysis of gene expression in tetraploid neurons.

In this chapter, we describe the protocol routinely performed in our laboratory to reliably quantify the proportion of tetraploid neurons in specific neural tissues as well as a method for the purification of diploid and tetraploid nuclei to perform further analyses with their genomic DNA.

2 Materials

2.1 Mice and Zebrafish

Postnatal day 15 (P15) mice (C57BL/6J background) and 20-month-old zebrafish (WIK line) were used in this study to obtain cerebral hemispheres and rhombencephalic derivatives, respectively. These tissues were frozen and stored at -80°C .

2.2 Oligonucleotides

Primers used in this chapter are those named as “inside forward” (TAGAGAGTTTATAAAGTTAG) and “inside reverse” (ACTAA AACAAAAACAACA) in the study by Li et al. [69]. For pyrosequencing, the inside forward primer was biotinylated at the 5' end, and the pyrosequencing primer was TAATACAACAACAACAA TAACAATC. PCR reactions were performed as described in [69].

2.3 Reagents

- Triton X-100 (Sigma-Aldrich).
- Protease inhibitor cocktail tablets (cOmplete, Mini, EDTA-free) (Roche).
- Bovine serum albumin (BSA) (stock 30 mg/mL in PBS) (Sigma-Aldrich).
- Fetal calf serum (Life Technologies).
- Mouse Anti-NeuN antibody (clone A60) (EMD Millipore).
- Rabbit monoclonal Anti-NeuN antibody [clone A60] (Abcam).
- Alexa Fluor 488 Goat Anti-Mouse IgG (H + L) Antibody (Life Technologies).
- Alexa Fluor 488 Donkey Anti-Rabbit IgG (H + L) Antibody (Life Technologies).
- Propidium iodide (stock: 1 mg/mL) (Sigma-Aldrich).
- DAPI (stock 200 $\mu\text{g/mL}$) (Sigma-Aldrich).

- RNase A (10 mg/mL, inactivated by boiling as indicated by the manufacturer) (Sigma-Aldrich).
- PyroMark Gold Q26 reagents (Qiagen).
- DNeasy Blood & Tissue Kit (Qiagen).
- EpiTect Bisulfite Kit (Qiagen).
- DNA Clean & Concentrator-5 Kit (ZymoResearch).
- pGEM-T Easy (Promega).
- PureYield Plasmid Miniprep System (Promega).

2.4 Buffers

- Phosphate-buffered saline (PBS).
- Nuclear isolation buffer (NIB): DNase-free PBS containing 0.1% Triton X-100 and 1× complete Protease Inhibitor Cocktail (Roche).

2.5 Equipment

- 15 mL tube (Falcon).
- 1.5 mL Standard and DNA LoBind tubes (Eppendorf).
- 0.5 mL Protein LoBind tubes (Eppendorf).
- 7 mL Dounce homogenizer (WHEATON).
- Epifluorescence E80i Nikon microscope equipped with a DXM 1200 digital camera.
- Centrifuge 5415 R (Eppendorf).
- Autoclaved 40 µm nylon filters.
- FACSAria cytometer (BD Biosciences) equipped with a 488 nm Coherent Sapphire solid-state and 633 nm JDS Uniphase HeNe air-cooled laser [emission filters: BP 530/30 (FITC), BP 616/23 (PE-Texas Red) and BP 660/20 (APC) for Alexa 488, PI, and Alexa 647, respectively].
- PSQ96MA pyrosequencer (Biotage).

2.6 Software

- FACSDiva (BD Biosciences).
- Pyro Q-CpG software (Biotage).
- QUMA software (<http://quma.cdb.riken.jp>).

2.7 Data Analysis and Statistics

Sanger sequencing results were analyzed for change in methylation pattern with the quantification tool for methylation analysis (QUMA) [70]. Data are shown as mean \pm s.e.m. For pyrosequencing, statistics were calculated over all CpGs with passed quality. Statistical significance was determined using two-tailed Student's *t* test.

3 Methods

3.1 Flow Cytometric Characterization of Tetraploid Neurons in Vertebrates: General Considerations

The development of a flow cytometric procedure for the characterization of both cells and organelles based on size and/or fluorescent labeling has been a major scientific breakthrough in biological sciences. This technology was initially developed for the analysis of isolated cells that can easily be passed through a laser beam. This includes unicellular organisms [71], dissociated cell lines [72], or biological samples containing single cells including blood [73], sperm [74], gynecologic samples [75], or bladder washings [76]. Nevertheless, further developments of the technique made flow cytometry also useful for the analysis and isolation of cells or cell organelles derived from compact tissues, provided that they are previously dissociated through enzymatic and/or mechanic procedures. This was not an exception for the brain, and protocols for the analysis of DNA content were initially described already in the late 1970s, using formalin-fixed cell nuclei obtained by tissue grinding from solid brain tumors [77].

We routinely use flow cytometry for DNA quantification in isolated cell nuclei of neural origin and in living cells [33, 36, 57, 64, 65, 68, 78]. In this section, we describe the procedure developed by our laboratory for the quantification of tetraploid neurons in vertebrate neural tissues using unfixed cell nuclei immunolabeled with antibodies specific for nucleus-located neuronal markers [36]. The use of unfixed cell nuclei results in low coefficients of variation, thus facilitating the identification of low-abundance nuclear populations. By using this procedure, we have been able to detect a small proportion of tetraploid neurons in several neural structures from different species, including human (unpublished results), mouse [33, 36], chick [33, 57], and zebrafish (see below).

Our method is based on tissue grinding with a Dounce homogenizer followed by two centrifugation steps to first remove undissociated tissue and then to isolate fresh cell nuclei. These cell nuclei are subsequently immunolabeled with an anti-NeuN antibody to identify those that are derived from neurons. This method has the advantage of allowing the characterization of any neuronal type provided that it can be identified by a nuclear marker. We have used this method to show that the major population of tetraploid neurons in the mouse cerebral cortex express CTIP2 [36], a known marker of subcortical projection neurons located in layers V and VI [79]. This observation, together with the finding that retinal tetraploid neurons are mainly a subpopulation of retinal ganglion cells, indicates that neuronal tetraploidy is mainly associated with projection neurons. The use of antibodies against the early-response genes *c-Fos* and *Erg-1*, known to be expressed in active neurons, also allowed us to demonstrate that tetraploid neurons are functional [36].

Along this section we explain our protocol for the identification and quantification of tetraploidy in neurons using the

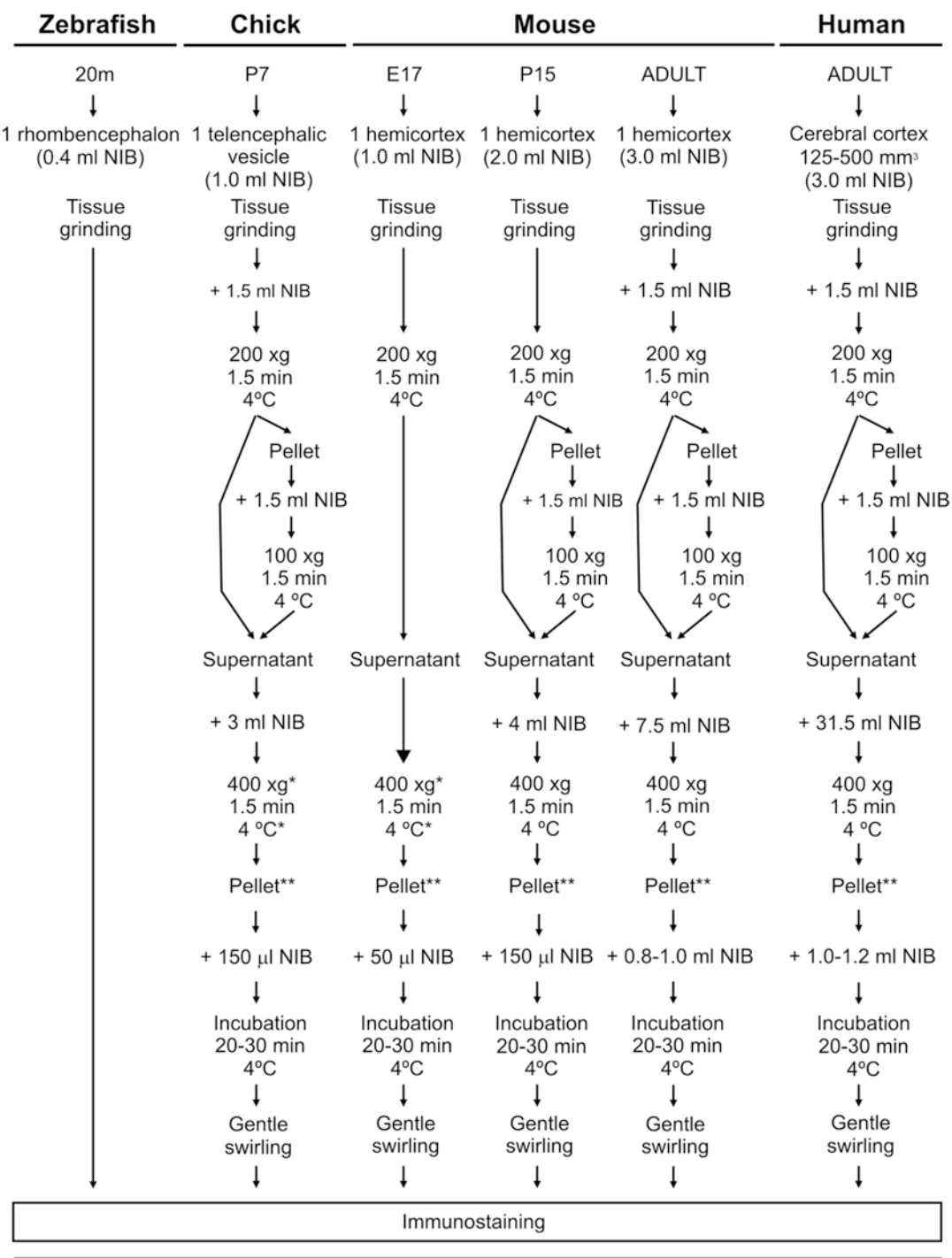
rhombencephalic derivatives of the zebrafish brain as a model system. In addition, we describe an optimized protocol for the isolation by FACS of cell nuclei from diploid and tetraploid neurons obtained from P15 cerebral cortex. Then, we describe how to obtain genomic DNA from these neurons and, finally, we present an example of the genomic analysis that can be performed in these neurons. For this aim we have chosen the analysis of CpG methylation in genomic imprinting (i.e., monoallelic expression of genes depending on parental origin). Most imprinted genes are located in clusters distributed throughout the whole genome [80], and they are co-regulated by cis-acting imprinting control elements that contain chromosome-specific DMRs. Epigenetic methylation of the CpGs constituting these DMRs is established in the germline and then this methylation pattern is tightly maintained in all somatic tissues during adulthood [81, 82]. In this chapter we describe how to study the methylation pattern of the *Rasgrf1*-imprinted domain [69] in diploid versus tetraploid cortical neurons from P15 mice.

3.2 Analysis and Quantification of Neuronal Tetraploidy in Zebrafish Neural Tissues

Our analyses using freshly isolated cell nuclei have mainly been focused on the mouse cerebral cortex, but they can also be applied to other neural structures such as the hippocampus and the olfactory bulb in this species. In addition, we have succeeded in the characterization of tetraploid neurons in brain tissue from other species including zebrafish and chick as well as from postmortem human cerebral tissue using minor modifications of the basic protocol (see Fig. 2). In this section, we describe the protocol for the characterization of tetraploid neurons using the neural structures derived from the rhombencephalon in zebrafish as a model system. Whenever required, indications of the necessary modifications for other species are described in the Notes section.

3.2.1 Preparation of Nuclear Suspensions and Immunolabeling

There are a number of variables to consider when cell nuclei suspensions are being prepared, including the species, age, and neural tissue under analysis. In the case at hands, experiments are performed with the rhombencephalon-derived tissue from one 20-month-old zebrafish (one neural structure per sample), previously snap-frozen on dry ice and stored at -80°C . The tissue is placed in 0.4 mL of ice-cold, nuclear isolation buffer (NIB) (see Note 1), and cell nuclei are then released by mechanical disaggregation using a Dounce homogenizer (with sequential use of both loose and tight pestles). The characteristics of zebrafish brain allow a complete mechanical dissociation and the obtention of very clear nuclear suspensions, but removal of small amounts of undissociated tissue could be necessary for other tissues using a centrifugation step (see Note 2). In some instances, small debris present in the supernatants obtained after this centrifugation step can be removed through subsequent washes with NBT (see Note 3).



*This step can be eliminated if the supernatant lacks cell debris
**Confirm absence of nuclei in supernatant through DAPI staining

Fig. 2 Flow diagram describing the protocol for the isolation of fresh cell nuclei from the indicated tissues. These cell nuclei can be used for immunostaining with neuronal markers with nuclear location and flow cytometry analysis

Once the nuclear preparation is obtained, cell nuclei are immunostained with antibodies against specific nuclear markers (NeuN in this case). For this purpose, we recommend to use commercially available antibodies known to work for chromatin immunoprecipitation as they are likely to recognize the native structure of the antigen. An improvement of our methodology is that both primary and secondary antibodies are added together to the cellular preparations (optimal dilutions have to be empirically deduced). This avoids unnecessary washing steps that may result in bias or generalized loss of cell nuclei. This makes our protocol highly accurate and yielding representative results, something that is crucial for this type of analysis. As a first approximation, primary antibodies should be diluted around 0.5–1.0 times the recommended dilution for immunohistochemistry. Secondary antibodies are usually employed at 1/500 dilution (*see Note 4*). For the example illustrated in Fig. 3, an anti-NeuN rabbit monoclonal antibody was used at 1/600 dilution and an Alexa 488 anti-rabbit antibody was diluted 500 times. Immunostaining is performed in an Eppendorf tube containing the isolated unfixed nuclei described above plus 5% fetal calf serum (FCS) and 1.25 mg/mL bovine serum albumin (BSA) (final concentrations). In this way, labeling is specific as blocking proteins are co-added with the antibodies. It is important to prepare control samples in which the primary antibodies are excluded. When the availability of the samples is limited, this control can be obtained by taking small aliquots from the experimental points, which are then joined together. Nuclear suspensions with antibodies are finally incubated O/N at 4 °C in the dark.

3.2.2 Flow Cytometric Analysis

NeuN-immunostained nuclei are filtered through a 40 µm nylon filter, and the volume adjusted to 0.5 mL (*see Note 5*) with PBS containing PI and DNase-free RNase I at a final concentration of 40 and 25 µg/mL, respectively, and then incubated for 30 min at room temperature (RT). The quality of the nuclei and specificity of immunostaining signal have to be checked by fluorescence microscopy (Fig. 3a).

In our experiments, the analysis by flow cytometry is performed with a FACSAria cytometer. Nevertheless, any flow cytometer available in the market would be equally efficient for this kind of analysis. Immunostained samples are passed through the cytometer and data are collected. In our case, we use Alexa 488 to label NeuN and PI for DNA quantification. Therefore, the analysis is performed with the BP 530/30 and BP 616/23 emission filters (used for Alexa 488 and PI, respectively). Data are then analyzed with the FACSDiva package. To this aim, gating is adjusted to remove cellular debris, which is easily differentiated from nuclei due to its inability to incorporate PI. Aggregates are then removed after plotting DNA pulse area versus its corresponding pulse height following the procedure described in [55] (Fig. 3b). Then, the

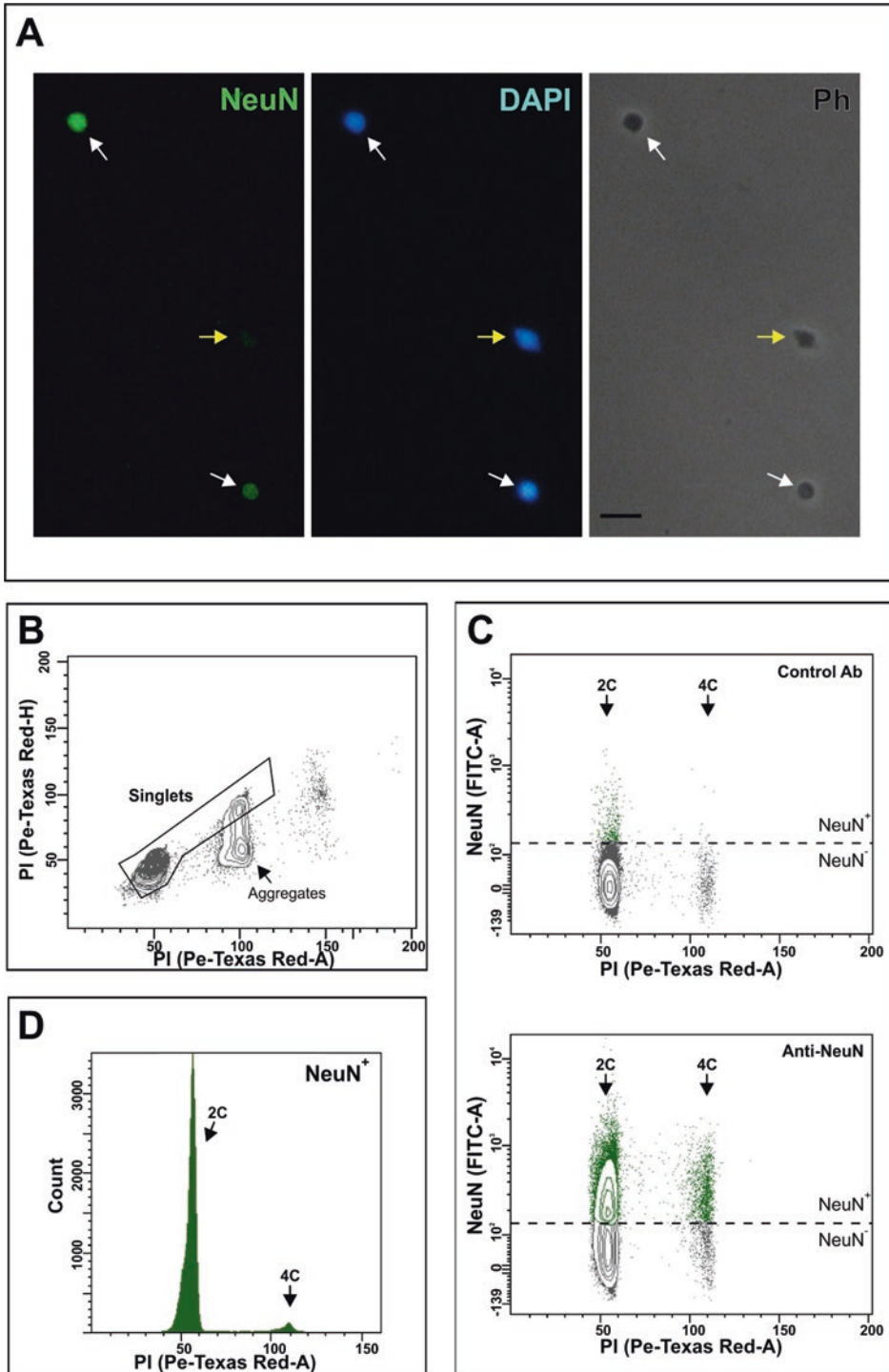


Fig. 3 Electronic gating method for zebrafish cell nuclei immunostained with NeuN and analyzed by flow cytometry. (a) NeuN-positive nuclei (*white arrows*) and NeuN-negative nucleus (*yellow arrow*) labeled with DAPI are shown. Ph: phase contrast. Bar: 10 μ m. (b) Fresh cell nuclei were isolated from neural tissues derived from the rhombencephalon of 20 m zebrafish, stained with propidium iodide (PI), and subjected to flow

NeuN-positive nuclear population is selected and plotted against PI intensity to quantify the proportion of neuronal tetraploid and diploid nuclei (Fig. 3c). A histogram is finally created for the quantification of the tetraploid neuron population. Figure 3d illustrates a representative result obtained from this analysis, indicating that $4.21 \pm 0.41\%$ (mean \pm s.e.m.) of neurons are tetraploid in the neural tissues derived from the rhombencephalon in the adult zebrafish ($n = 8$).

3.3 Isolation of Diploid and Tetraploid Neurons from Vertebrate Neural Tissues to Perform Epigenetic Studies

Besides the common use of flow cytometry for the characterization and quantification of tetraploid neurons in specific tissues, FACS technology can also be used for the isolation of genomic DNA from these neurons (i.e., 4C NeuN-positive cell nuclei). To this aim, we usually employ a nozzle with a 100 μ m diameter, a size that does not perturb nuclear integrity. It is important to point out that cell sorting is a challenging approach when the abundance of the target population is low, as it is the case for tetraploid neurons. A single round of cell sorting is able to enrich the tetraploid nuclear population to around 20% as a maximum. Therefore, two rounds of cell sorting are required to enrich the tetraploid population to 80–90%. As a consequence, the final number of tetraploid nuclei that are obtained is quite low. Another reason for a reduced yield of tetraploid neuronal nuclei is the necessity of checking its proportion within the enriched sample. This analysis results in the loss of a significant number of isolated nuclei, which are passed through the flow cytometer following the protocol described above. As an average, we usually recover around 1000–9000 tetraploid nuclei out of four P15 mouse hemicortices, and about 100–500 cell nuclei are used for this validation step. The rest can be used for genomic DNA extraction. The isolation of basically pure diploid neurons is easier, as this population represents around 98–99% of total neurons and around half of all cell nuclei in the cerebral cortex. Therefore, a single round of FACS is sufficient for the isolation of highly enriched diploid neuron populations. We usually isolate genomic DNA from 200,000 diploid cell nuclei.

3.3.1 Isolation of Cell Nuclei from Diploid and Tetraploid Cortical Neurons

For the isolation of cortical neurons of diploid and tetraploid DNA content, one cortical hemisphere from P15 mice is homogenized using a Dounce homogenizer in 2 mL of ice-cold NIB (see Fig. 2). Undissociated tissue is then removed by centrifugation at $200 \times g$

Fig. 3 (continued) cytometric analysis. After gating on forward scattering area vs. side scattering area (not shown), DNA content was Fig. 3 (continued) assessed by plotting PI (Pe-Texas Red-H) vs. PI (Pe-Texas Red-A) levels. Diploid and tetraploid nuclei were subsequently gated from these plots (singlets), while the doublets of diploid nuclei were discarded (aggregates). **(c)** DNA content from the gated cell population containing both diploid (2C) and tetraploid (4C) events, plotted against Neu immunolabeling levels (FITC-A), is shown. A control analysis performed with nuclei labeled with secondary antibodies is also shown. **(d)** A histogram showing the count numbers of diploid (2C) and tetraploid (4C) nuclei is shown

for 1.5 min (4 °C) as described above. The supernatant is then diluted with 4 mL of NIB, while the pellet is resuspended with 1.5 mL NIB and centrifuged at $100 \times g$ for 1.5 min at 4 °C, and the resulting supernatant is added to the previous one (see Fig. 2). The mixed supernatant is then distributed in five Eppendorf tubes and centrifuged at $400 \times g$ to isolate cell nuclei and remove the cell debris from the sample. Each pellet is then incubated for 30 min with 30 μ L NIB on ice, resuspended by gently swirling, joined into one single Eppendorf tube (150 μ L in total), and directly used for immunolabeling. To this aim, the cell nuclei are incubated O/N in the dark at 4 °C with a mouse monoclonal antibody against NeuN diluted 1/1000 and a secondary Alexa 488 anti-mouse antibody used at a 1/500 dilution in NIB containing 5% fetal calf serum (FCS) and 1.25 mg/mL bovine serum albumin (BSA). In parallel, control samples are prepared lacking the primary antibody. After the incubation period, immunostained nuclei are filtered through a 40 μ m nylon filter and, then, DNA staining is performed with 40 μ g/mL PI, while RNA is removed with 25 μ g/mL DNase-free RNase I, which is incubated for 30 min at RT. After analyzing the quality of the nuclei and the specificity of the immunostaining by fluorescence microscopy, cell nuclei are sorted through a FACS Aria cytometer. Gating is set as described above to isolate diploid and tetraploid NeuN-positive nuclei on DNA LoBind and Protein LoBind Eppendorf tubes, respectively. Diploid neuronal cell nuclei are selected for further analysis while the tetraploid enriched fraction is subjected to a second round of cell sorting to isolate tetraploid neuronal cell nuclei. In this case, tetraploid events are collected in DNA LoBind tubes.

3.3.2 Genomic DNA Isolation

To obtain genomic DNA from diploid and tetraploid neurons, the isolated cell nucleus suspensions are centrifuged at $20,000 \times g$ for 5 min, and then most of the supernatant is removed (just 200 μ L is kept). The genomic DNA is then isolated from this pellet by using the DNeasy Blood & Tissue Kit, a procedure that requires no phenol extraction and involves minimal handling. To this aim, 20 μ L of the proteinase K solution and 200 μ L of buffer AL (both included in the kit) are added to the remnant volume and then the mix is vortexed to assure cell nuclear lysis. This buffer allows selective binding of DNA to the DNeasy membrane located within a spin column also included in the kit. After two wash steps, DNA is eluted with 40 μ L of the buffer AE included in the kit. This volume is optimal for DNA yields below 1 μ g.

3.3.3 Bisulfite Treatment of Genomic DNA from Diploid and Tetraploid Neurons

The genomic DNA isolated in the previous step can be used for several types of analysis. In this chapter we used CpG methylation analysis in imprinting. The most common technique for the analysis of CpG methylation is the treatment of genomic DNA with

bisulfite followed by sequencing [83]. Bisulfite converts cytosine residues to uracil, but leaves 5-methylcytosine residues unaffected. Therefore, only methylated cytosines are retained when genomic DNA has been treated with bisulfite.

For this analysis, 20 μL of the genomic DNA obtained in the previous step is bisulfite-converted using EpiTect Bisulfite Kit. To this aim, the DNA solution is combined with 85 μL of bisulfite mix solution and 35 μL of DNA protect buffer, and then the protocol is performed as indicated in the specifications of the kit.

3.3.4 Analysis of CpG Methylation Levels in Imprinting Control Regions of Diploid and Tetraploid Neurons through Sanger Sequencing

The Sanger sequencing procedure is the classical method for the analysis of CpG methylation [83]. For this analysis, a representative DNA sequence within the DMR of interest is PCR amplified from the bisulfite-converted DNA using specific oligonucleotides designed for the bisulfite-converted sequence (i.e., cytosines substituted by thymines). As a general rule, whenever possible the presence of GpGs should be avoided from the oligonucleotide sequences. In case this cannot be applied, degenerate oligonucleotides recognizing both cytidines and thymines should be used. In addition, when the CpG methylation levels are studied by pyrosequencing (see below) one of the oligonucleotides should be biotinylated in its 5' end. If possible, primers should be designed flanking a single-nucleotide polymorphism (SNP), which is used to determine the parental origin of the specific strand when analyzed by Sanger sequencing. In this case, the analyzed tissue should be derived from mice obtained by crossing of two genetic backgrounds carrying the mentioned SNP. Finally, the analysis of the sequencing data is performed with the QUMA software [70], using its standard settings.

To study the CpG methylation levels of the DMR controlling the imprinted *Rasgrf1* region [69] (Fig. 4a) in diploid and tetraploid neurons, two primers flanking this DMR (see Fig. 4b) are used. These primers are specific for the bisulfite-converted sequence, and they encompass 29 CpG sites within the *Rasgrf1*-imprinted domain. PCR amplification is then performed using the PCR conditions described in [69], and the amplification product is cleaned with the DNA Clean & Concentrator-5 kit. After PCR product subcloning into the pGEM-Teasy vector, a significant number of positive clones are subjected to miniprep preparation of plasmids with the PureYield Plasmid Miniprep System, and these plasmids are subsequently sequenced through the Sanger method. Figure 5 illustrates the result obtained from this analysis, which indicates that at P15 both diploid and tetraploid neurons seem to show more than the expected 50% of methylated CpGs.

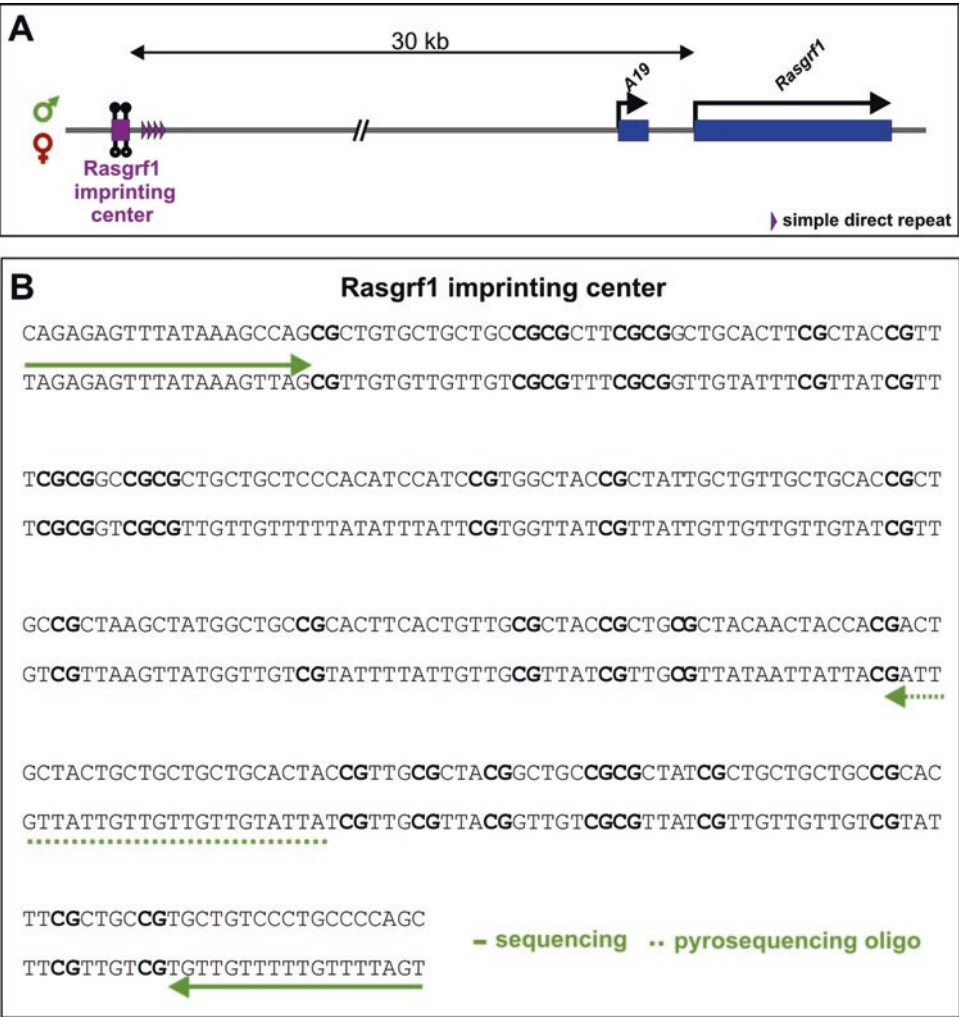


Fig. 4 Scheme showing the genomic structure of murine *Rasgrf1*-imprinted domain and the sequence analyzed within the *Rasgrf1*-imprinting center. **(a)** Genes constituting the *Rasgrf1*-imprinting domain are shown. These genes show paternal allele-specific expression. The *Rasgrf1*-imprinting center is 30 kb upstream of these genes. Methylated CpG sites from this imprinting center are represented by *solid circles*, while *open circles* represent non-methylated CpGs from the *Rasgrf1*-controlling DMR. **(b)** Genomic DNA sequence corresponding to the region analyzed in this chapter. *Solid arrows* : primers used for amplification (for pyrosequencing, the forward primer was 5'-biotinylated). *Dashed arrow* : primer used for pyrosequencing

3.3.5 Analysis of CpG Methylation Levels in Imprinting Control Regions of Diploid and Tetraploid Neurons through Pyrosequencing

Sanger sequencing depends on random incorporation of selective dideoxynucleotides lacking a 3'-OH group required for the DNA extension reaction, thus leading to products of different lengths that can be assigned to a particular sequence [84]. In contrast to this method, which needs a relatively limited number of clones that are then sequenced, pyrosequencing has a superior quantitative capacity based on the detection of the levels of pyrophosphate that is released when a deoxyribonucleotide triphosphate is added to

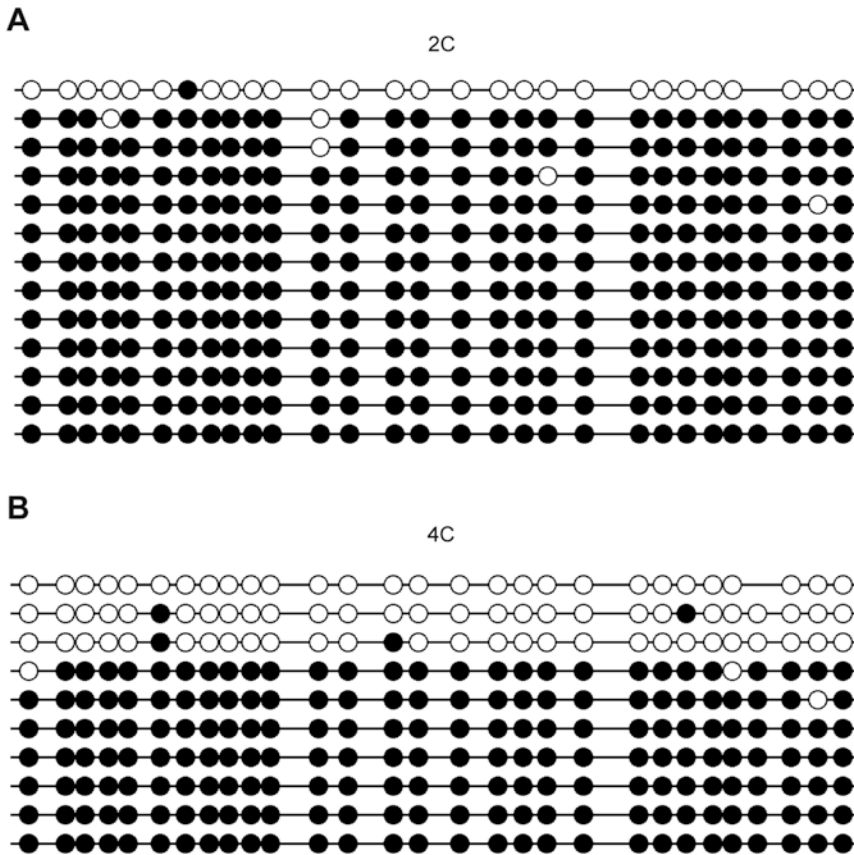


Fig. 5 Sanger sequencing methylation analysis of the *Rasgrf1*-imprinting control region shown in Fig. 4. Results were analyzed for change in methylation pattern with the QUMA software, which generates a panel of *solid* and *open circles* indicating methylated and unmethylated CpGs, respectively

the end of a nascent strand of DNA [85]. Although this technique can be used to reveal unknown DNA sequences, it is not as efficient as other available sequencing procedures. Nevertheless, pyrosequencing has proved to be highly efficient and sensitive for the analysis of single-nucleotide mutations in short DNA sequences (usually less than 80 nucleotides in length) [86]. Deoxyribonucleotide triphosphates are added to the reaction following the expected sequence. This results in the release of pyrophosphate (PPi), whose concentration can be monitored as a peak of light emission. This allows the quantification of the mutations that are present in the DNA sequence (see scheme in Fig. 6). This procedure is therefore quite convenient for the quantification of the proportion of cytosine conversion to thymine in bisulfite-treated DNA. The proportion of CpG methylation present within the DMRs of an imprinting DNA domain is therefore easily quantified by this procedure when DNA fragments initially enriched in CpGs are generated by PCR amplification from bisulfite-converted genomic DNA.

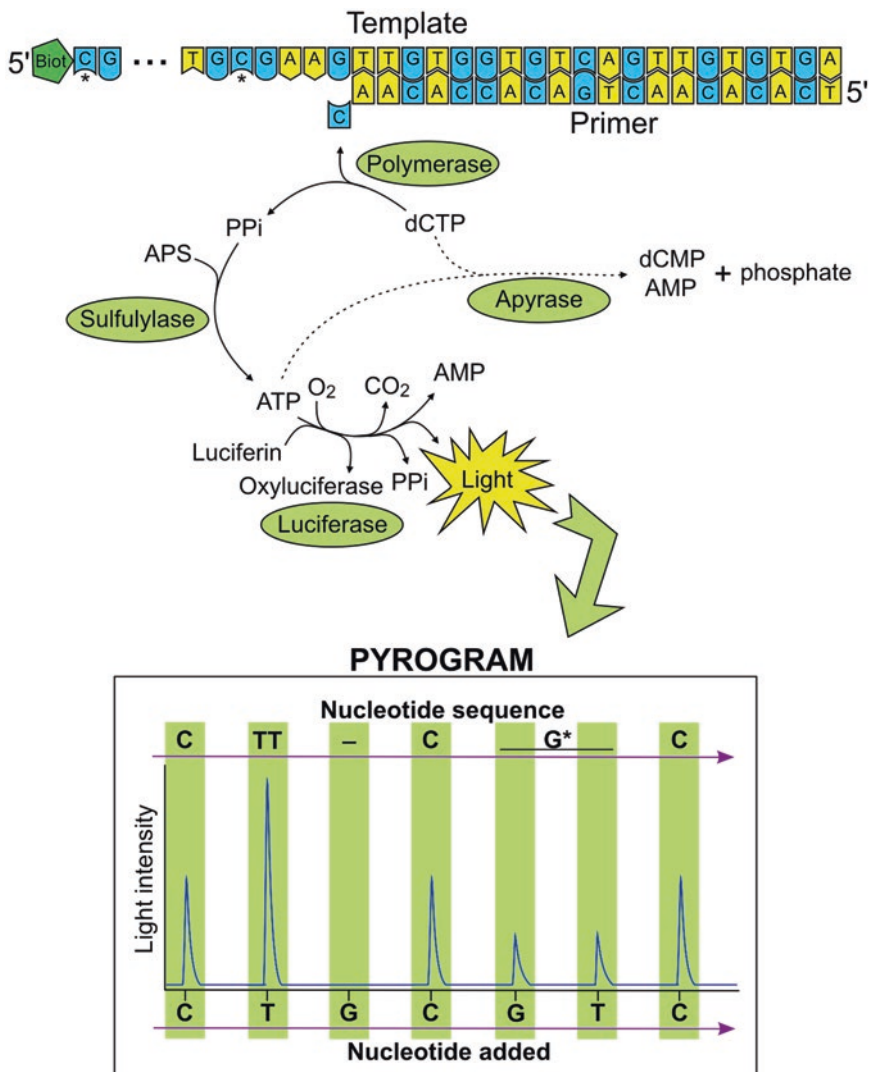


Fig. 6 Scheme illustrating the pyrosequencing method. A PCR amplicon generated with a pair of primers, one of them biotinylated at its 5' end, is denatured, and then the biotinylated strand is immobilized through its biotinylated 5' end. Asterisks indicate CpGs that may be methylated: if C is not methylated, C should be substituted by T (in the complementary strand, G should be substituted by A). This template is hybridized to a sequencing primer and then incubated with DNA polymerase, ATP sulfurylase, and luciferase together with the substrates adenosine 5' phosphosulfate (APS) and luciferin. The addition of the deoxynucleoside triphosphate complementary to the one present in the template sequence (dCTP in this example) initiates the pyrosequencing reaction. dATP α S is used instead of dATP to avoid noise, since the former is not a substrate for a luciferase. The incorporation by DNA polymerase of the correct, complementary dNTP releases pyrophosphate (PPi). Then, ATP sulfurylase converts PPi to ATP in the presence of APS, and this ATP acts as a substrate for luciferase, which converts luciferin to oxyluciferin and generates light in proportion to the amount of ATP. This light is detected by a camera and shown as a pyrogram. Finally, the presence of apyrase results in the degradation of the unincorporated nucleotides and ATP, and the reaction can restart with another nucleotide. As an internal control of the pyrosequencing method, a reaction step can be performed with a noncomplementary nucleotide (dGTP in this example), thus resulting in lack of light signal. When the template contains more than one identical nucleotide (the pair TT in the example), the peak of light is proportional to the number of identical nucleotides present in the sequence. The percentage of methylated CpGs is calculated after successive addition of dGTP and dTTP (in the example 50% of the CpGs are methylated since the peak of light is similar for both dGTP and dTTP)

To confirm the observation that the *Rasgrfl* DMR contains a high proportion of methylated CpGs, bisulfite-treated DNA is PCR amplified using the same primers as above, the forward primer being biotinylated at the 5' end. The PCR product is then bound to streptavidin-coated Sepharose beads (GE-Healthcare) through the biotin tag and denatured to generate single-stranded DNA to allow annealing of an internal sequencing primer (16 pmol of each per reaction) (see Materials). Pyrosequencing is then performed with a PSQTM96MA instrument using 25 μ L of amplified DNA product and PyroMark Gold Q96 reagents. Pyrosequencing data analysis is finally done using the Pyro Q-CpG software. This technique can only resolve an average of 80 nucleotides upstream from the sequencing primer. Therefore, only five CpGs could be quantified with this technique (Fig. 7). This analysis indicated that $78.31 \pm 3.88\%$ ($n = 4$) of CpGs were methylated in diploid neurons whereas this proportion was $75.07 \pm 4.77\%$ ($n = 3$) in tetraploid neurons (Fig. 7). This result, which indicates that tetraploid neurons show similar proportion of methylated CpGs in the DMR of the *Rasgrfl*-imprinting domain, is consistent with the results obtained through Sanger sequencing illustrated in Fig. 5. For some unknown reason, this domain seems to be overmethylated in P15 cortical neurons.

4 Notes

1. This volume should be increased to 1 mL for one telencephalic hemisphere from P7 chick or E17 mouse embryos, to 2 mL for one hemicortex of P15 mice, and to 3 mL for either one hemicortex of adult mouse or one cube of 5–8 mm edge of human brain (see Fig. 2).
2. The standard procedure for the removal of the undissociated tissue (mainly fibers and blood vessels) is by centrifugation ($200 \times g$ for 1.5 min at 4 °C in 1.5 mL minifuge tubes) discarding pellet. If tissue from either adult mouse or human brain is processed, 1.5 mL of cold NIB should be added prior to the centrifugation step (see Fig. 2). Additionally, when quantification or high recovery is needed, the pellet of this centrifugation step can be resuspended with 1.5 mL of cold NIB, and then centrifuged at $100 \times g$ for 1.5 min at 4 °C to remove any dissociated cell nuclei that may remain among the undissociated tissue. The supernatant is then added to the previous one (see Fig. 2 for details).
3. For postnatal and adult mouse tissues the supernatant is further diluted with NIB (see Fig. 2 for details). In all cases (except for zebrafish rhombencephalon-derived neural structures) tissues are centrifuged at $400 \times g$ for 4 min at 4 °C.

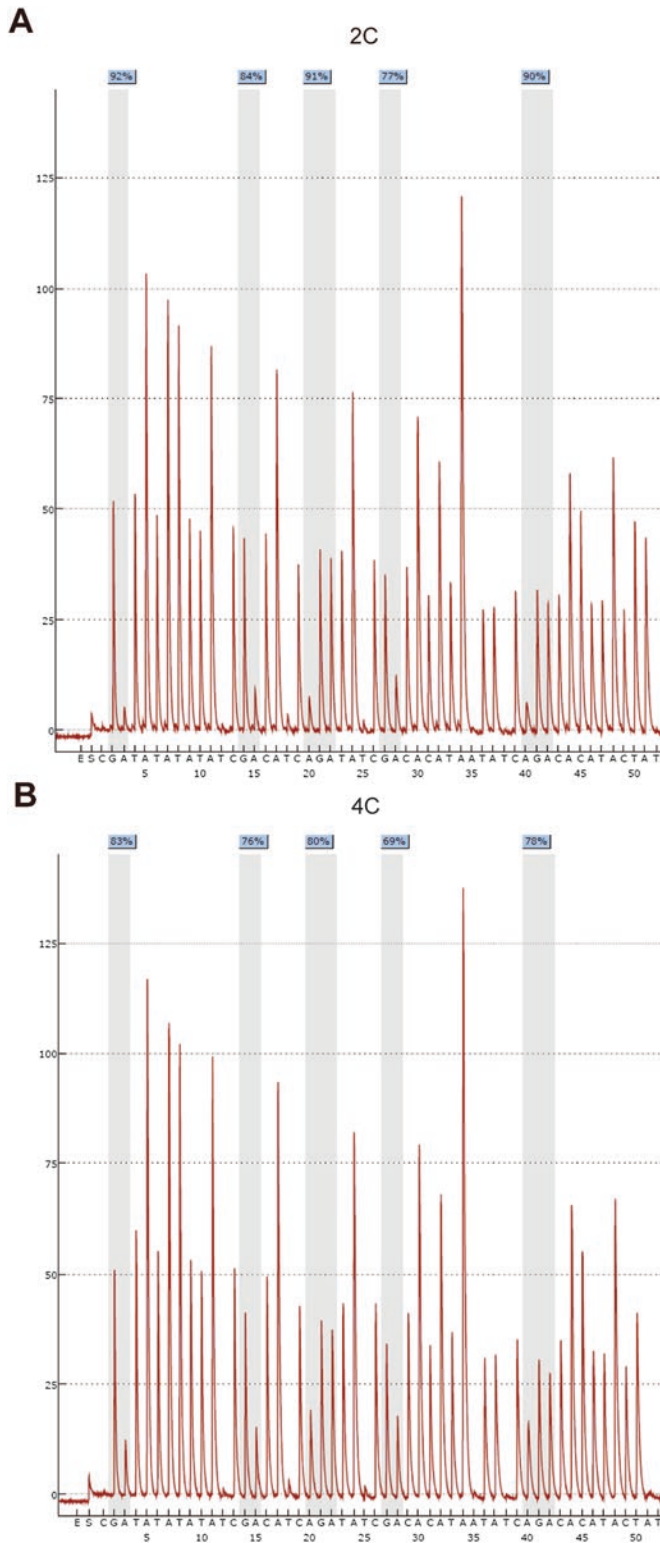


Fig. 7 Representative pyrograms from diploid (a) and tetraploid (b) neuronal cell nuclei, corresponding to the *Rasgrf1*-imprinting control region shown in Fig. 4, are shown. Blue boxes: Percentage of methylation in any particular CpG analyzed. This value is calculated as the ratio between the G peak (complementary to the bisulfite-protected ^{Met}C) and the G + A peaks (complementary, respectively, to the bisulfite-protected ^{Met}C plus the bisulfite-unprotected C)

Supernatants containing cellular debris are then discarded and the pellet is then incubated at 4 °C in cold NIB for at least 1 h, before mechanical resuspension by gentle swirl of the vial. The volume for this resuspension step differs from tissue to tissue (see Fig. 2 for details). A quality control that is required at this stage is the analysis under the microscope of a small aliquot of the isolated nuclei stained with 100 ng/mL DAPI. This analysis informs about the quality and purity of the nuclear preparation. It is convenient to use only cell nuclear preparations in which cell debris are kept to a minimum.

4. Fluorophores are chosen depending on the flow cytometer. It is imperative that both absorption and emission spectra of the fluorophores are compatible with the laser and filters of cytometer and that their emission spectra do not overlap.
5. This volume can be adjusted to 1–3 mL in other tissues to avoid high-concentrated cell nucleus suspensions.

Acknowledgments

We thank Gonzalo García de Polavieja and Aixa Morales for the zebrafish specimens used in this study. This research was supported by the Ministerio de Economía y Competitividad (grant numbers SAF2015-68488-R) and a R&D contract with Tetraneuron S.L.

References

1. Masterson J (1994) Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* 264:421–424
2. Sugimoto-Shirasu K, Roberts K (2003) "big it up": endoreduplication and cell-size control in plants. *Curr Opin Plant Biol* 6:544–553
3. Madlung A, Wendel JF (2013) Genetic and epigenetic aspects of polyploid evolution in plants. *Cytogenet Genome Res* 140:270–285
4. Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annu Rev Genet* 34:401–437
5. Li S, Yin L, Cole ES, Udani RA, Karrer KM (2006) Progeny of germ line knockouts of *ASI2*, a gene encoding a putative signal transduction receptor in *Tetrahymena Thermophila*, fail to make the transition from sexual reproduction to vegetative growth. *Dev Biol* 295:633–646
6. Kolics B, Ács Z, Chobanov DP, Orci KM, Qiang LS, Kovács B, Kondorosy E, Decsi K, Taller J, Specziár A, Orbán L, Müller T (2012) Re-visiting phylogenetic and taxonomic relationships in the genus *Saga* (Insecta: Orthoptera). *PLoS One* 7:e42229
7. Leggatt RA, Iwama GK (2003) Occurrence of polyploidy in the fishes. *Rev Fish Biol Fish* 13:237–246
8. Vernon JA, Butsch J (1957) Effect of tetraploidy on learning and retention in the salamander. *Science* 125:1033–1034
9. Szaro BG, Tompkins R (1987) Effect of tetraploidy on dendritic branching in neurons and glial cells of the frog, *Xenopus laevis*. *J Comp Neurol* 258:304–316
10. Trifonov VA, Paoletti A, Caputo Barucchi V, Kalinina T, O'Brien PC, Ferguson-Smith MA, Giovannotti M (2015) Comparative chromosome painting and NOR distribution suggest a complex hybrid origin of triploid *Lepidodactylus lugubris* (Gekkonidae). *PLoS One* 10:e0132380
11. Frawley LE, Orr-Weaver TL (2015) Polyploidy. *Curr Biol* 25:R353–R358
12. Biesterfeld S, Gerres K, Fischer-Wein G, Böcking A (1994) Polyploidy in non-neoplastic tissues. *J Clin Pathol* 47:38–42
13. Edgar BA, Orr-Weaver TL (2001) Endoreplication cell cycles: more for less. *Cell* 105:297–306

14. Ullah Z, Lee CY, Depamphilis ML (2009) Cip/kip cyclin-dependent protein kinase inhibitors and the road to polyploidy. *Cell Div* 4:10
15. Ullah Z, Lee CY, Lilly MA, DePamphilis ML (2009) Developmentally programmed endoreplication in animals. *Cell Cycle* 8:1501–1509
16. Pandit SK, Westendorp B, de Bruin A (2013) Physiological significance of polyploidization in mammalian cells. *Trends Cell Biol* 23:556–566
17. Unhavaithaya Y, Orr-Weaver TL (2012) Polyploidization of glia in neural development links tissue growth to blood-brain barrier integrity. *Genes Dev* 26:31–36
18. Lilly MA, Spradling AC (1996) The drosophila endocycle is controlled by cyclin E and lacks a checkpoint ensuring S-phase completion. *Genes Dev* 10:2514–2526
19. Buntrock L, Marec F, Krueger S, Traut W (2012) Organ growth without cell division: somatic polyploidy in a moth, *Ephestia kuehniella*. *Genome* 55:755–763
20. Vitrat N, Cohen-Solal K, Pique C, Le Couedic JP, Norol F, Larsen AK, Katz A, Vainchenker W, Debili N (1998) Endomitosis of human megakaryocytes are due to abortive mitosis. *Blood* 91:3711–3723
21. Ullah Z, Kohn MJ, Yagi R, Vassilev LT, DePamphilis ML (2008) Differentiation of trophoblast stem cells into giant cells is triggered by p57/Kip2 inhibition of CDK1 activity. *Genes Dev* 22:3024–3036
22. Vinogradov AE, Anatskaya OV, Kudryavtsev BN (2001) Relationship of hepatocyte ploidy levels with body size and growth rate in mammals. *Genome* 44:350–360
23. Zanet J, Freije A, Ruiz M, Coulon V, Sanz JR, Chiesa J, Gandarillas A (2010) A mitosis block links active cell cycle with human epidermal differentiation and results in endoreplication. *PLoS One* 5:e15701
24. Nagata Y, Jones MR, Nguyen HG, McCrann DJ, St Hilaire C, Schreiber BM, Hashimoto A, Inagaki M, Earnshaw WC, Todokoro K, Ravid K (2005) Vascular smooth muscle cell polyploidization involves changes in chromosome passenger proteins and an endomitotic cell cycle. *Exp Cell Res* 305:277–291
25. Kellerman S, Moore JA, Zierhut W, Zimmer HG, Campbell J, Gerdes AM (1992) Nuclear DNA content and nucleation patterns in rat cardiac myocytes from different models of cardiac hypertrophy. *J Mol Cell Cardiol* 24:497–505
26. Coggeshall RE, Yaksta BA, Swartz FJ (1970) A cytophotometric analysis of the DNA in the nucleus of the giant cell, R-2, in *Aplysia*. *Chromosoma* 32:205–212
27. Manfredi Romanini MG, Fraschini A, Bernocchi G (1973) DNA content and nuclear area in the neurons of the cerebral ganglion in *Helix pomatia*. *Ann Histochem* 18:49–58
28. Castelfranco AM, Hartline DK (2016) Evolution of rapid nerve conduction. *Brain Res* 1641(Pt a):11–33
29. Lasek RJ, Dower WJ (1971) *Aplysia californica*: analysis of nuclear DNA in individual nuclei of giant neurons. *Science* 172:278–280
30. Swift H (1953) Quantitative aspects of nuclear nucleoproteins. *Int Rev Cytol* 2:1–76
31. Lapham LW (1968) Tetraploid DNA content of Purkinje neurons of human cerebellar cortex. *Science* 159:310–312
32. Mosch B, Morawski M, Mittag A, Lenz D, Tarnok A, Arendt T (2007) Aneuploidy and DNA replication in the normal human brain and Alzheimer's disease. *J Neurosci* 27:6859–6867
33. Morillo SM, Escoll P, de la Hera A, Frade JM (2010) Somatic tetraploidy in specific chick retinal ganglion cells induced by nerve growth factor. *Proc Natl Acad Sci U S A* 107:109–114
34. Shirazi Fard S, Jarrin M, Boije H, Fillon V, All-Eriksson C, Hallböök F (2013) Heterogenic final cell cycle by chicken retinal Lim1 horizontal progenitor cells leads to heteroploid cells with a remaining replicated genome. *PLoS One* 8:e59133
35. Shirazi Fard S, All-Eriksson C, Hallböök F (2014) The heterogenic final cell cycle of chicken retinal Lim1 horizontal cells is not regulated by the DNA damage response pathway. *Cell Cycle* 13:408–417
36. López-Sánchez N, Frade JM (2013) Genetic evidence for p75^{NTR}-dependent tetraploidy in cortical projection neurons from adult mice. *J Neurosci* 33:7488–7500
37. Ovejero-Benito MC, Frade JM (2013) Brain-derived neurotrophic factor-dependent cdk1 inhibition prevents G2/M progression in differentiating tetraploid neurons. *PLoS One* 8:e64890
38. Ovejero-Benito MC, Frade JM (2015) p27^{Kip1} participates in the regulation of endoreplication in differentiating chick retinal ganglion cells. *Cell Cycle* 14:2311–2322
39. Kingsbury MA, Yung YC, Peterson SE, Westra JW, Chun J (2006) Aneuploidy in the normal and diseased brain. *Cell Mol Life Sci* 63:2626–2641
40. Ying QL, Nichols J, Evans EP, Smith AG (2002) Changing potency by spontaneous fusion. *Nature* 416:545–548

41. Feulgen R, Rossenbeck H (1924) Mikroskopisch-chemischer Nachweis einer Nukleinsäure von Typus der Thymonukleinsäure und die darauf beruhende selektive Färbung von Zellkernen in mikroskopischen Präparaten. *Hoppe Seylers Z Physiol Chem* 135:203–248
42. Chieco P, Derenzini M (1999) The Feulgen reaction 75 years on. *Histochem Cell Biol* 111:345–358
43. Woodard J, Gorovsky M, Swift H (1966) DNA content of a chromosome of *Trillium erectum*: effect of cold treatment. *Science* 151:215–216
44. Herman CJ, Lapham LW (1968) DNA content of neurons in the cat hippocampus. *Science* 160:537
45. Herman CJ, Lapham LW (1969) Neuronal polyploidy and nuclear volumes in the cat central nervous system. *Brain Res* 15:35–48
46. Museridze DP, Svanidze IK, Macharashvili DN (1975) Content of DNA and dry weight of the nuclei of neurons of the external geniculate body and retina of the eye in guinea pigs. *Sov J Dev Biol* 5:269–272
47. Swartz FJ, Bhatnagar KP (1981) Are CNS neurons polyploid? A critical analysis based upon cytophotometric study of the DNA content of cerebellar and olfactory bulb neurons of the bat. *Brain Res* 208:267–281
48. Mosch B, Mittag A, Lenz D, Arendt T, Tárnok A (2006) Laser scanning cytometry in human brain slices. *Cytometry A* 69:135–138
49. Fulwyler MJ (1965) Electronic separation of biological cells by volume. *Science* 150:910–911
50. Kamensky LA, Melamed MR, Derman H (1965) Spectrophotometer: new instrument for ultrarapid cell analysis. *Science* 150:630–631
51. Van Dilla MA, Trujillo TT, Mullaney PF, Coulter JR (1969) Cell microfluorometry: a method for rapid fluorescence measurement. *Science* 163:1213–1214
52. Hulett HR, Bonner WA, Barrett J, Herzenberg LA (1969) Cell sorting: automated separation of mammalian cells as a function of intracellular fluorescence. *Science* 166:747–749
53. Hoshino T, Nomura K, Wilson CB, Knebel KD, Gray JW (1978) The distribution of nuclear DNA from human brain-tumor cells. *J Neurosurg* 49:13–21
54. Bernocchi G, Barni S (1985) On the heterogeneity of Purkinje neurons in vertebrates. Cytochemical and morphological studies of chromatin during eel (*Anguilla Anguilla* L.) life cycle. *J Hirnforsch* 26:227–235
55. Nunez R (2001) DNA measurement and cell cycle analysis by flow cytometry. *Curr Issues Mol Biol* 3:67–70
56. Martinez-Morales PL, Quiroga AC, Barbas JA, Morales AV (2010) SOX5 controls cell cycle progression in neural progenitors by interfering with the WNT-beta-catenin pathway. *EMBO Rep* 11:466–472
57. López-Sánchez N, Ovejero-Benito MC, Borreguero L, Frade JM (2011) Control of neuronal ploidy during vertebrate development. *Results Probl Cell Differ* 53:547–563
58. Hasbold J, Hodgkin PD (2000) Flow cytometric cell division tracking using nuclei. *Cytometry* 40(3):230–237
59. Westra JW, Barral S, Chun J (2009) A reevaluation of tetraploidy in the Alzheimer's disease brain. *Neurodegener Dis* 6:221–229
60. Westra JW, Rivera RR, Bushman DM, Yung YC, Peterson SE, Barral S, Chun J (2010) Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. *J Comp Neurol* 518:3981–4000
61. Nüsse M, Beisker W, Hoffmann C, Tarnok A (1990) Flow cytometric analysis of G1- and G2/M-phase subpopulations in mammalian cell nuclei using side scatter and DNA content measurements. *Cytometry* 11:813–821
62. Darzynkiewicz Z, Traganos F, Kapuscinski J, Staiano-Coico L, Melamed MR (1984) Accessibility of DNA in situ to various fluorochromes: relationship to chromatin changes during erythroid differentiation of friend leukemia cells. *Cytometry* 5:355–363
63. Darzynkiewicz Z (2011) Critical aspects in analysis of cellular DNA content. In: Robinson JP (ed) *Current protocols in cytometry*. Wiley, New York
64. López-Sánchez N, Frade JM (2013) Cell cycle analysis in the vertebrate brain using immunolabeled fresh cell nuclei. *Bio-protocol* 3:e973
65. López-Sánchez N, Frade JM (2015) Flow cytometric analysis of DNA synthesis and apoptosis in central nervous system using fresh cell nuclei. *Methods Mol Biol* 1254:33–42
66. Shapiro HM (2003) *Practical flow cytometry*, 4th edn. Wiley, New York
67. Murciano A, Zamora J, López-Sánchez J, Frade JM (2002) Interkinetic nuclear movement may provide spatial clues to the regulation of neurogenesis. *Mol Cell Neurosci* 21:285–300
68. Slaninová I, López-Sánchez N, Šebrová K, Vymazal O, Frade JM, Táborská E (2016) Introduction of macarpine as a novel cell-permeant DNA dye for live cell imaging and flow cytometry sorting. *Biol Cell* 108:1–18
69. Li JY, Lees-Murdock DJ, GL X, Walsh CP (2004) Timing of establishment of paternal

- methylation imprints in the mouse. *Genomics* 84:952–960
70. Kumaki Y, Oda M, Okano M (2008) QUMA: quantification tool for methylation analysis. *Nucleic Acids Res* 36(suppl 2):W170–W175
 71. Hutter KJ, Eipel HE (1978) Flow cytometric determinations of cellular substances in algae, bacteria, moulds and yeasts. *Antonie Van Leeuwenhoek* 44:269–282
 72. Gray JW, Carver JH, George YS, Mendelsohn ML (1977) Rapid cell cycle analysis by measurement of the radioactivity per cell in a narrow window in S phase (RCSI). *Cell Tissue Kinet* 10:97–109
 73. Wilder ME, Cram LS (1977) Differential fluorescence of human lymphocytes as measured by flow cytometry. *J Histochem Cytochem* 25:888–891
 74. Van Dilla MA, Gledhill BL, Lake S, Dean PN, Gray JW, Kachel V, Barlogie B, Göhde W (1977) Measurement of mammalian sperm deoxyribonucleic acid by flow cytometry. Problems and approaches. *J Histochem Cytochem* 25:763–773
 75. Jensen RH (1977) Chromomycin A3 as a fluorescent probe for flow cytometry of human gynecologic samples. *J Histochem Cytochem* 25:573–579
 76. Pedersen T, Larsen JK, Krarup T (1978) Characterization of bladder tumours by flow cytometry on bladder washings. *Eur Urol* 4:351–355
 77. Frederiksen P, Reske-Nielsen E, Bichel P (1978) Flow cytometry in tumours of the brain. *Acta Neuropathol* 41:179–183
 78. Morillo SM, Abanto EP, Román MJ, Frade JM (2012) Nerve growth factor-induced cell cycle reentry in newborn neurons is triggered by p38^{MAPK}-dependent E2F4 phosphorylation. *Mol Cell Biol* 32:2722–2737
 79. Arlotta P, Molyneaux BJ, Chen J, Inoue J, Kominami R, Macklis JD (2005) Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron* 45:207–221
 80. Wan LB, Bartolomei MS (2008) Regulation of imprinting in clusters: noncoding RNAs versus insulators. *Adv Genet* 61:207–223
 81. Leonhardt H, Page AW, Weier HU, Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71:865–873
 82. Constância M, Pickard B, Kelsey G, Reik W (1998) Imprinting mechanisms. *Genome Res* 8:881–900
 83. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* 89:1827–1831
 84. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–5467
 85. Ronaghi M, Karamohamed S, Pettersson B, Uhlén M, Nyrén P (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem* 242:84–89
 86. Langae T, Ronaghi M (2005) Genetic variation analyses by pyrosequencing. *Mutat Res* 573:96–102